

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:34:18 ; Search time 150.545 Seconds  
(without alignments)  
70.046 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLHNQEVTAEDGTQAEKSKPCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A\_Geneseq\_21:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

9: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	100.0	24	4	AAB68627
2	132	100.0	47	4	AAB68665
3	128	97.0	24	4	AAB68663
4	128	97.0	387	9	ADZ47819 Amino aci
5	128	97.0	402	7	ADM29354
6	128	97.0	419	4	AAE09203
7	128	97.0	419	4	AAE09211
8	128	97.0	419	4	AAE09210
9	128	97.0	419	4	AAE09205
10	128	97.0	419	4	AAE09208
11	128	97.0	419	4	AAE09212
12	128	97.0	419	4	AAE09206
13	128	97.0	419	4	AAE09207
14	128	97.0	419	4	AAE09213
15	128	97.0	419	4	AAE09202
16	128	97.0	419	4	AAE09214
17	128	97.0	419	4	AAE09212
18	128	97.0	419	4	AAE09215
19	128	97.0	419	4	AAE09216
20	128	97.0	419	4	AAE09181
21	128	97.0	419	4	AAE09200
22	128	97.0	419	4	AAE09204
23	128	97.0	419	4	AAE09209
24	128	97.0	419	5	AAE20348

98 128 97.0 1255 7 ADF45048 Adf45048 Human kin  
99 128 97.0 1255 8 ADH13187 Adh13187 Human mal  
100 128 97.0 1255 8 ADJ66554 Adj66554 Her2 prot

## ALIGNMENTS

RESULT 1  
AAB68627  
ID AAB68627 standard; peptide; 24 AA.  
XX AC AAB68627;  
XX DT 30-APR-2001 (first entry)  
XX XX  
DE HER-2 B cell epitope #6.  
XX  
KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;  
KW ovarian; lung; prostate; colon.  
XX  
OS Homo sapiens.  
XX  
PN WO200108636-A2.  
XX  
PD 08-FEB-2001.  
XX  
PF 03-AUG-2000; 2000WO-US021222.  
XX  
PR 03-AUG-1999; 99US-0146869P.  
XX  
XX (OHIS ) UNIV OHIO STATE.  
XX  
XX Kaumaya PT, Stevens VC, Triozzi PL;  
PI WPI; 2001-182849/18.  
XX  
XX Compositions comprising polypeptides and polynucleotides for stimulating  
PT the immune system and for treating malignancies associated with  
PT overexpression of the HER-2 protein.

XX Claim 1; Page 37; 5lpp; English.  
XX  
CC The present invention relates to compositions for stimulating the immune  
CC system and for treating malignancies associated with overexpression of  
CC the HER-2 protein. The compositions comprise immunogenic groups of the  
CC HER-2 proteins. The present sequence is one such peptide used in the  
CC compositions of the present invention. The compositions can be used for  
CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers  
XX  
SQ Sequence 24 AA;  
Query Match 100.0%; Score 132; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 PLHNQEVTAEDGTQRAEKCKPCA 24

RESULT 2  
AAB68665  
ID AAB68665 standard; peptide; 47 AA.  
XX AC AAB68665;  
XX DT 30-APR-2001 (first entry)  
XX XX  
DE HER-2 B cell epitope #15.  
XX  
KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;  
KW ovarian; lung; prostate; colon.

XX Unidentified.  
OS  
XX WO200108636-A2.  
PN  
XX  
PD 08-FEB-2001.  
XX  
PF 03-AUG-2000; 2000WO-US021222.  
XX  
PR 03-AUG-1999; 99US-0146869P.  
XX  
XX (OHIS ) UNIV OHIO STATE.  
XX  
XX Kaumaya PT, Stevens VC, Triozzi PL;  
PI WPI; 2001-182849/18.  
XX  
XX Compositions comprising polypeptides and polynucleotides for stimulating  
PT the immune system and for treating malignancies associated with  
PT overexpression of the HER-2 protein.  
XX  
XX Example 9; Page 31; 5lpp; English.  
XX  
CC The present invention relates to compositions for stimulating the immune  
CC system and for treating malignancies associated with overexpression of  
CC the HER-2 protein. The compositions comprise immunogenic groups of the  
CC HER-2 proteins. The present sequence is one such peptide used in the  
CC compositions of the present invention. The compositions can be used for  
CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers  
XX  
SQ Sequence 47 AA;  
Query Match 100.0%; Score 132; DB 4; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.6e-12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 PLHNQEVTAEDGTQRAEKCKPCA 24

RESULT 3  
AAB68663  
ID AAB68663 standard; peptide; 24 AA.  
XX AC AAB68663;  
XX DT 30-APR-2001 (first entry)  
XX  
XX HER-2 B cell epitope #13.  
XX  
XX Cytostatic; immune response; HER-2; human; epitope; cancer; breast;  
XX ovarian; lung; prostate; colon.  
XX  
XX Unidentified.  
OS  
XX WO200108636-A2.  
PN  
XX  
PD 08-FEB-2001.  
XX  
PF 03-AUG-2000; 2000WO-US021222.  
XX  
PR 03-AUG-1999; 99US-0146869P.  
XX  
XX (OHIS ) UNIV OHIO STATE.  
XX  
XX Kaumaya PT, Stevens VC, Triozzi PL;  
PI WPI; 2001-182849/18.  
XX  
XX Compositions comprising polypeptides and polynucleotides for stimulating  
PT the immune system and for treating malignancies associated with  
PT overexpression of the HER-2 protein.

XX PS Example 8; Page 30; 51pp; English.  
XX CC The present invention relates to compositions for stimulating the immune  
XX CC system and for treating malignancies associated with overexpression of  
XX CC the HER-2 protein. The compositions comprise immunogenic groups of the  
XX CC HER-2 proteins. The present sequence is one such peptide used in the  
XX CC compositions of the present invention. The compositions can be used for  
XX CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers  
XX SQ Sequence 24 AA;  
  
Query Match 97.0%; Score 128; DB 4; Length 24;  
Best Local Similarity 95.8%; Pred. No. 3e-12;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PLHNQEVTAEDGTQRAEKCKSKPCA 24  
Db 1 PLHNQEVTAEDGTQRAEKCKSKPCA 24  
|||||  
  
RESULT 4  
AD247819  
ID AD247819 standard; protein; 387 AA.  
XX AC AD247819;  
XX DT 30-JUN-2005 (first entry)  
XX DE Amino acid sequence for human ErbB-2 variant VI.  
XX KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
XX KW prognosis; cancer; cytostatic.  
XX OS Homo sapiens.  
XX PN WO2005033133-A2.  
XX PD 14-APR-2005.  
XX PF 04-OCT-2004; 2004WO-US030903.  
XX PR 03-OCT-2003; 2003US-0507953P.  
XX PR 10-NOV-2003; 2003US-0518321P.  
XX PR 09-AUG-2004; 2004US-0599583P.  
XX PR 07-SEP-2004; 2004US-0607326P.  
XX (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.  
XX PI Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX WPI; 2005-285403/29.  
DR N-PSDB; AD247818.  
XX  
XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
PT cancer.  
XX  
XX Claim 42; SEQ ID NO 26; 246pp; English.  
XX  
XX The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
XX cancer. This sequence represents human ErbB-2 variant VI.  
XX

SQ Sequence 387 AA;  
  
Query Match 97.0%; Score 128; DB 9; Length 387;  
Best Local Similarity 95.8%; Pred. No. 6.7e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PLHNQEVTAEDGTQRAEKCKSKPCA 24  
Db 316 PLHNQEVTAEDGTQRAEKCKSKPCA 339  
|||||  
  
RESULT 5  
ADM29354  
ID ADM29354 standard; protein; 402 AA.  
XX AC ADM29354;  
XX DT 20-MAY-2004 (first entry)  
XX DE Human novel protein NOV20b.  
XX KW human; novel protein; NOV; cancer; immune associated disorder.  
XX OS Homo sapiens.  
XX PN WO2003064628-A2.  
XX PD 07-AUG-2003.  
XX PF 03-FEB-2003; 2003WO-US003401.  
XX PR 01-FEB-2002; 2002US-0353287P.  
XX PR 01-FEB-2002; 2002US-0353301P.  
XX PR 12-FEB-2002; 2002US-0356371P.  
XX PR 12-FEB-2002; 2002US-0356424P.  
XX PR 13-FEB-2002; 2002US-0356531P.  
XX PR 20-FEB-2002; 2002US-0358239P.  
XX PR 26-FEB-2002; 2002US-0359603P.  
XX PR 27-FEB-2002; 2002US-0359848P.  
XX PR 27-FEB-2002; 2002US-0359860P.  
XX PR 15-MAR-2002; 2002US-0365049P.  
XX PR 22-MAR-2002; 2002US-0366802P.  
XX PR 17-MAY-2002; 2002US-0381666P.  
XX PR 18-JUN-2002; 2002US-0389531P.  
XX PR 19-JUN-2002; 2002US-038910P.  
XX PR 25-JUN-2002; 2002US-0391516P.  
XX PR 02-JUL-2002; 2002US-0393265P.  
XX PR 07-AUG-2002; 2002US-0401825P.  
XX PR 09-AUG-2002; 2002US-0402395P.  
XX PR 12-AUG-2002; 2002US-0402867P.  
XX PR 23-AUG-2002; 2002US-0405401P.  
XX PR 23-AUG-2002; 2002US-0405820P.  
XX (CURA-) CURAGEN CORP.  
PA  
XX Alsobrook JP, Bader JS, Berghs C, Burgess CE, Casman SJ;  
XX Catterton E, Chaudhuri A, Edinger SR, Ellerman K, Gerlach VL;  
PI Gorman L, Guo X, Herrmann JL, Ji W, Khrantsov NV, Li L, Miller CE;  
PI Ort T, Patturajan M, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;  
PI Spytek KA, Vernet CAM, Zhong H, Zhong M;  
XX WPI; 2003-646149/61.  
DR N-PSDB; ADM29353.  
XX  
XX New NOVX polypeptide, useful for the manufacture of a medicament for  
PT treating e.g., cancer or immune associated disorders.  
XX  
XX Claim 1; SEQ ID NO 98; 606pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of novel  
CC human proteins (NOV proteins). The DNA and protein sequences of the  
CC invention are useful for the manufacture of a medicament for treating a  
XX syndrome associated with a human disease comprising a pathology  
CC

CC associated with the protein, such as: cancer or immune associated  
 CC disorders. The present amino acid sequence represents a NOV protein of  
 XX the invention.  
 XX

SQ Sequence 402 AA;

Query Match 97.0%; Score 128; DB 7; Length 402;  
 Best Local Similarity 95.8%; Pred. NO. 7e-11;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
 Db 316 PLHNQEVTAEDGTQRCCKPCA 339

RESULT 6

AAE09203  
 ID AAE09203 standard; protein; 419 AA.

AC AAE09203;

XX 15 NOV-2001 (first entry)

DE Human p68HER-2 generic protein variant 1.

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIA; variant.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Region 1..340

FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"

FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"

FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown  
 in the specification"

FT Domain 341..419  
 FT /label= ECDIIIA variant

FT Misc-difference 342 /note= "Extracellular domain IIIa variant"

FT /note= "p68HER-2 generic sequence (AAE09181) Xaa  
 substituted with Ser"

FT Misc-difference 345 /label= Unknown

FT Misc-difference 346 /label= Unknown

FT Misc-difference 356 /label= Unknown

FT Misc-difference 358 /label= Unknown

FT Misc-difference 361 /label= Unknown

FT Misc-difference 376 /label= Unknown

FT Misc-difference 394 /label= Unknown

FT Misc-difference 404 /label= Unknown

FT Misc-difference 413 /label= Unknown

XX WO200161356-A1.

XX 23-AUG-2001.

XX 16-FEB-2001; 2001WO-US005327.

XX 16-FEB-2000; 2000US-00506079.

XX

PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX Clinton G, Henner WD, Evans A;  
 XX WPI; 2001-529934/58.

XX New polypeptide, which binds to the extracellular domain of HER-2 for the  
 treatment of hard tumors.

XX Example 11; Page; 61pp; English.

CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 particularly a polypeptide that binds to the extracellular domain (ECD)  
 of HER-2 at a site that is different from the binding site of humanised  
 antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 invention is based upon the initial discovery of an alternative HER-2  
 mRNA transcript with 274 bp insert of intron 8. The translation product  
 of the alternative transcript is a truncated HER-2 protein designated  
 p68HER-2 which lacks the transmembrane and intracellular domains of  
 p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 nucleic acids encoding these are useful to treat, diagnose and identify  
 solid tumours. The present sequence is human p68HER-2 generic protein  
 containing ECDIIIA variant sequence. Note: The present sequence is not  
 shown in the specification but is derived from p68HER-2 generic sequence  
 (SEQ ID NO:2) shown in the sequence listing (AAE09181)

XX Sequence 419 AA;

Query Match 97.0%; Score 128; DB 4; Length 419;  
 Best Local Similarity 95.8%; Pred. NO. 7.4e-11;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24

Db 316 PLHNQEVTAEDGTQRCCKPCA 339

RESULT 7

AAE09211

ID AAE09211 standard; protein; 419 AA.

XX AAE09211;

XX 15-NOV-2001 (first entry)

DE Human p68HER-2 generic protein variant 9.

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIA; variant.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Region 1..340

FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"

FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"

FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown  
 in the specification"

FT Domain 341..419

FT /label= ECDIIIA variant

FT /note= "Extracellular domain IIIa variant"

FT Misc-difference 342 /label= Unknown

FT Misc-difference 345 /label= Unknown

FT Misc-difference 346 /label= Unknown

FT



FT Misc-difference 356 /label= Unknown  
 FT Misc-difference 358 /label= Unknown  
 FT Misc-difference 361 /label= Unknown  
 FT Misc-difference 376 /label= Unknown  
 FT Misc-difference 394 /label= Unknown  
 FT Misc-difference 404 /label= Unknown  
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa  
 substituted with Leu"  
 FT Misc-difference 413 /label= Unknown  
 FT  
 FT  
 PN WO200161356-A1.  
 XX  
 XX  
 PD 23-AUG-2001.  
 XX  
 XX 16-FEB-2001; 2001WO-US005327.  
 XX  
 XX 16-FEB-2000; 2000US-00506079.  
 XX  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 XX Clinton G, Henner WD, Evans A;  
 XX  
 XX WPI; 2001-529934/58.  
 DR  
 XX  
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the  
 PT treatment of hard tumors.  
 XX  
 XX Example 11; Page; 61pp; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIA variant sequence. Note: The present sequence is not  
 CC shown in the specification but is derived from p68HER-2 generic sequence  
 CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)  
 XX  
 SQ Sequence 419 AA;  
  
 Query Match 97.0%; Score 128; DB 4; Length 419;  
 Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 PLHNQVTAEDGTQRAEKCCKPCA 24  
 Db 316 PLHNQVTAEDGTQRAEKCCKPCA 339  
  
 RESULT 8  
 AAE09210  
 ID AAE09210 standard; protein; 419 AA.  
 AC  
 AC AAE09210;  
 XX  
 XX 15-NOV-2001 (first entry)  
 DT  
 XX Human p68HER-2 generic protein variant 8.  
 DE  
 XX

KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIA; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the parent sequence shown  
 in the specification"  
 FT Domain 341..419  
 FT /label= ECDIIIA\_variant  
 FT /note= "Extracellular domain IIIA variant"  
 FT Misc-difference 342  
 FT /label= Unknown  
 FT Misc-difference 345  
 FT /label= Unknown  
 FT Misc-difference 346  
 FT /label= Unknown  
 FT Misc-difference 356  
 FT /label= Unknown  
 FT Misc-difference 358  
 FT /label= Unknown  
 FT Misc-difference 361  
 FT /label= Unknown  
 FT Misc-difference 376  
 FT /label= Unknown  
 FT Misc-difference 394  
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa  
 substituted with Arg"  
 FT Misc-difference 404  
 FT /label= Unknown  
 FT Misc-difference 413  
 FT /label= Unknown  
 XX WO200161356-A1.  
 XX  
 XX 23-AUG-2001.  
 XX  
 XX 16-FEB-2001; 2001WO-US005327.  
 PF  
 XX 16-FEB-2000; 2000US-00506079.  
 PR  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 XX Clinton G, Henner WD, Evans A;  
 XX  
 XX WPI; 2001-529934/58.  
 DR  
 XX  
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the  
 PT treatment of hard tumors.  
 XX  
 XX Example 11; Page; 61pp; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIA variant sequence. Note: The present sequence is not  
 CC shown in the specification but is derived from p68HER-2 generic sequence  
 CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)  
 XX  
 SQ Sequence 419 AA;  
  
 Query Match 97.0%; Score 128; DB 4; Length 419;  
 Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 PLHNQVTAEDGTQRAEKCCKPCA 24  
 Db 316 PLHNQVTAEDGTQRAEKCCKPCA 339  
  
 RESULT 8  
 AAE09210  
 ID AAE09210 standard; protein; 419 AA.  
 AC  
 AC AAE09210;  
 XX  
 XX 15-NOV-2001 (first entry)  
 DT  
 XX Human p68HER-2 generic protein variant 8.  
 DE  
 XX

```
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;
Query Match 97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PLHNOEVTAE DGTORAEKCKSPCA 24
Db 316 PLHNOEVTAE DGTQRCCKSPCA 339
RESULT 9
AAE09205
ID AAE09205 standard; protein; 419 AA.
AC AAE09205;
XX
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 3.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= "ECDIIIA variant"
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
, PA
Clinton G, Henner WD, Evans A;
WPI; 2001-529934/58.
New polypeptide, which binds to the extracellular domain of HER-2 for the
treatment of hard tumors.
Example 11; Page; 61pp; English.
The invention relates to novel HER-2 (herstatin-2) antagonist
particularly a polypeptide that binds to the extracellular domain (ECD)
of HER-2 at a site that is different from the binding site of humanised
antibody, Herceptin, at an affinity of at least 10-8. The present
invention is based upon the initial discovery of an alternative HER-2
mRNA transcript with 274 bp insert of intron 8. The translation product
of the alternative transcript is a truncated HER-2 protein designated
p68HER-2 which lacks the transmembrane and intracellular domains of
p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
nucleic acids encoding these are useful to treat, diagnose and identify
solid tumours. The present sequence is human p68HER-2 generic protein
containing ECDIIIA variant sequence. Note: The present sequence is not
shown in the specification but is derived from p68HER-2 generic sequence
(SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;
Query Match 97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PLHNOEVTAE DGTORAEKCKSPCA 24
Db 316 PLHNOEVTAE DGTQRCCKSPCA 339
RESULT 10
AAE09208
ID AAE09208 standard; protein; 419 AA.
XX
AC AAE09208;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 6.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= "ECDIIIA variant"
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
, PA
```

```

FT      /label= Unknown
FT      Misc-difference 358
FT      /label= Unknown
FT      Misc-difference 361
FT      /label= Asp, Ala, Val
FT      Misc-difference 376
FT      /label= Unknown
FT      Misc-difference 394
FT      /label= Unknown
FT      Misc-difference 404
FT      /label= Unknown
FT      Misc-difference 413
FT      /label= Unknown
FT      WO200161356-A1.
FT      23-AUG-2001.
XX      16-FEB-2001; 2001WO-US005327.
XX      16-FEB-2000; 2000US-00506079.
XX      (UYOR-) UNIV OREGON HEALTH SCI.
XX      Clinton G, Henner WD, Evans A;
XX      WPI; 2001-529934/58.
XX      New polypeptide, which binds to the extracellular domain of HER-2 for the
XX      treatment of hard tumors.
XX      Example 11; Page; 61pp; English.
XX      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      of HER-2 at a site that is different from the binding site of humanised
XX      antibody, Herceptin, at an affinity of at least 108. The present
XX      invention is based upon the initial discovery of an alternative HER-2
XX      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      of the alternative transcript is a truncated HER-2 protein designated
XX      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      nucleic acids encoding these are useful to treat, diagnose and identify
XX      solid tumours. The present sequence is human p68HER-2 generic protein
XX      containing ECDIIIA variant sequence. Note: The present sequence is not
XX      shown in the specification but is derived from p68HER-2 generic sequence
XX      (SEQ ID NO:2) shown in the sequence listing (AAE09181).
XX      Sequence 419 AA;
XX      Query Match      97.0%; Score 128; DB 4; Length 419;
XX      Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX      Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX      OY      1 PLHNQVTAEDGTQRAECSKPCA 24
XX      Db      316 PLHNQVTAEDGTQRCSEKPCA 339
XX      RESULT 11
XX      AAE09206
XX      ID      AAE09206 standard; protein; 419 AA.
XX      AC      AAE09206;
XX      XX      15-NOV-2001 (first entry)
XX      DE      Human p68HER-2 generic protein variant 4.
XX      XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      KW      solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX      KW      (SEQ ID NO:2) shown in the sequence listing (AAE09181)

```

---

```

KW      ECDIIIA; variant.
XX      OS      Homo sapiens.
XX      FH      Key      Location/Qualifiers
XX      Region      1..340
XX      FT      Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"
XX      FT      /note= "Represented as Agn in the parent sequence shown
XX      FT      in the specification"
XX      FT      Misc-difference 125
XX      FT      /note= "Represented as Agn in the parent sequence shown
XX      FT      in the specification"
XX      FT      Domain      341..419
XX      FT      /label= ECDIIIA variant
XX      FT      /note= "Extracellular domain IIIa variant"
XX      FT      Misc-difference 342 /label= Unknown
XX      FT      Misc-difference 345 /label= Unknown
XX      FT      Misc-difference 346 /label= Unknown
XX      FT      Misc-difference 356 /label= Unknown
XX      FT      /note= "p68HER-2 generic sequence (AAE09181) Xaa
XX      FT      substituted with Gln"
XX      FT      Misc-difference 358 /label= Unknown
XX      FT      Misc-difference 361 /label= Unknown
XX      FT      Misc-difference 376 /label= Unknown
XX      FT      Misc-difference 394 /label= Unknown
XX      FT      Misc-difference 404 /label= Unknown
XX      FT      Misc-difference 413 /label= Unknown
XX      FT      /label= Unknown
XX      WO200161356-A1.
XX      23-AUG-2001.
XX      16-FEB-2001; 2001WO-US005327.
XX      16-FEB-2000; 2000US-00506079.
XX      (UYOR-) UNIV OREGON HEALTH SCI.
XX      Clinton G, Henner WD, Evans A;
XX      WPI; 2001-529934/58.
XX      New polypeptide, which binds to the extracellular domain of HER-2 for the
XX      treatment of hard tumors.
XX      Example 11; Page; 61pp; English.
XX      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      of HER-2 at a site that is different from the binding site of humanised
XX      antibody, Herceptin, at an affinity of at least 108. The present
XX      invention is based upon the initial discovery of an alternative HER-2
XX      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      of the alternative transcript is a truncated HER-2 protein designated
XX      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      nucleic acids encoding these are useful to treat, diagnose and identify
XX      solid tumours. The present sequence is human p68HER-2 generic protein
XX      containing ECDIIIA variant sequence. Note: The present sequence is not
XX      shown in the specification but is derived from p68HER-2 generic sequence
XX      (SEQ ID NO:2) shown in the sequence listing (AAE09181).
XX      Sequence 419 AA;
XX      Query Match      97.0%; Score 128; DB 4; Length 419;
XX      Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX      Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX      OY      1 PLHNQVTAEDGTQRAECSKPCA 24
XX      Db      316 PLHNQVTAEDGTQRCSEKPCA 339
XX      RESULT 11
XX      AAE09206
XX      ID      AAE09206 standard; protein; 419 AA.
XX      AC      AAE09206;
XX      XX      15-NOV-2001 (first entry)
XX      DE      Human p68HER-2 generic protein variant 4.
XX      XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      KW      solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX      KW      (SEQ ID NO:2) shown in the sequence listing (AAE09181)

```

```

XX SQ Sequence 419 AA;
XX
XX Query Match 97.0%; Score 128; DB 4; Length 419;
XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
Db 316 PLHNQEVTAEDGTQRCCKSPCA 339

RESULT 12
AAE09207
ID AAE09207 standard; protein; 419 AA.
XX
AC AAE09207;
XX
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 5.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
PI

```

---

```

XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 97.0%; Score 128; DB 4; Length 419;
XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
Db 316 PLHNQEVTAEDGTQRCCKSPCA 339

RESULT 13
AAE09213
ID AAE09213 standard; protein; 419 AA.
XX
AC AAE09213;
XX
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 11.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
PI

```

```

FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Asn"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 97.0%; Score 128; DB 4; Length 419;
XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 PLHNQEVTAEDGTORAEKCKPKCA 24
XX |||||||||||||||||||
XX Db 316 PLHNQEVTAEDGTORCEKCKPKCA 339
XX
XX RESULT 14
XX AAE09202
XX ID AAE09202 standard; protein; 419 AA.
XX
XX AC AAE09202;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant (Arg371Ile).
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;

```

```

KW ECDIIIA; variant.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIA variant"
FT
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 371 /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Arg
FT substituted with Ile"
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 97.0%; Score 128; DB 4; Length 419;
XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 PLHNQEVTAEDGTORAEKCKPKCA 24
XX |||||||||||||||||||
XX Db 316 PLHNQEVTAEDGTORCEKCKPKCA 339
XX
XX RESULT 14
XX AAE09202
XX ID AAE09202 standard; protein; 419 AA.
XX
XX AC AAE09202;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant (Arg371Ile).
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;

```

```
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match          97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRCCKSPCA 339

RESULT 15
AAE09214
ID AAE09214 standard; protein; 419 AA.
AC AAE09214;
DT 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 13.
DE
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
FH Key
FT Region 1..340
FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 125 /note= "Represented as Agn in the sequence shown in the
specification"
FT Domain 341..419
FT Misc-difference 342 /note= "Represented as Agn in the sequence shown in the
specification"
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa"
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 357 /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09183) Xaa
substituted with Cys"
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 371 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
FT /label= Unknown
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
```

```
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
treatment of hard tumors.
XX
XX Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 108. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:13) shown in the sequence listing (AAE09183)
XX
SQ Sequence 419 AA;

Query Match          97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRCCKSPCA 339

RESULT 16
AAE09212
ID AAE09212 standard; protein; 419 AA.
XX
AC AAE09212;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 10.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
FH Key
FT Region 1..340
FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
```

FT Misc-difference 346 /label= Unknown  
 FT 346 /label= Unknown  
 FT Misc-difference 356 /label= Unknown  
 FT 356 /label= Unknown  
 FT Misc-difference 358 /label= Unknown  
 FT 358 /label= Unknown  
 FT Misc-difference 361 /label= Unknown  
 FT 361 /label= Unknown  
 FT Misc-difference 376 /label= Unknown  
 FT 376 /label= Unknown  
 FT Misc-difference 394 /label= Unknown  
 FT 394 /label= Unknown  
 FT Misc-difference 404 /label= Unknown  
 FT 404 /label= Unknown  
 FT Misc-difference 413 /label= Unknown  
 FT 413 /label= Unknown  
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa  
 FT substituted with Asn"  
 FT  
 PN WO200161356-A1.  
 XX  
 XX 23-AUG-2001.  
 XX  
 XX 16-FEB-2001; 2001WO-US005327.  
 XX  
 XX 16-FEB-2000; 2000US-00506079.  
 XX  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 XX Clinton G, Henner WD, Evans A;  
 XX  
 XX WPI; 2001-529934/58.  
 XX  
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the  
 XX treatment of hard tumors.  
 XX  
 XX Example 11; Page; 61pp; English.  
 XX  
 XX The invention relates to novel HER-2 (herstatin-2) antagonist  
 XX particularly a polypeptide that binds to the extracellular domain (ECD)  
 XX of HER-2 at a site that is different from the binding site of humanised  
 XX antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 XX invention is based upon the initial discovery of an alternative HER-2  
 XX mRNA transcript with 274 bp insert of intron 8. The translation product  
 XX of the alternative transcript is a truncated HER-2 protein designated  
 XX p68HER-2 which lacks the transmembrane and intracellular domains of  
 XX p68HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 XX nucleic acids encoding these are useful to treat, diagnose and identify  
 XX solid tumours. The present sequence is human p68HER-2 generic protein  
 XX containing ECDIIIA variant sequence. Note: The present sequence is not  
 XX shown in the specification but is derived from p68HER-2 generic sequence  
 XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)  
 XX  
 XX SQ Sequence 419 AA;  
 XX  
 XX Query Match 97.08; Score 128; DB 4; Length 419;  
 XX Best Local Similarity 95.8; Pred. No. 7.4e-11;  
 XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX  
 OY 1 PLNQEVTAEDGTORAEKSKPCA 24  
 DB 316 PLNQEVTAEDGTORAEKSKPCA 339  
 XX  
 XX RESULT 17  
 XX AAE09215  
 XX ID AAE09215 standard; protein; 419 AA.  
 XX  
 XX AC AAE09215;  
 XX  
 XX DT 15-NOV-2001 (first entry)

XX DE Human p68HER-2 generic protein variant 14.  
 XX  
 XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIA; variant.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the sequence shown in the  
 FT specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the sequence shown in the  
 FT specification"  
 FT 341..419  
 FT Domain /label= ECDIIIA variant  
 FT /note= "Extracellular domain IIIa"  
 FT Misc-difference 342 /label= Unknown  
 FT Misc-difference 345 /label= Unknown  
 FT Misc-difference 346 /label= Unknown  
 FT Misc-difference 356 /label= Unknown  
 FT Misc-difference 357 /label= Unknown  
 FT Misc-difference 358 /label= Unknown  
 FT Misc-difference 361 /label= Unknown  
 FT Misc-difference 371 /label= Unknown  
 FT /note= "p68HER-2 generic sequence (AAE09183) Xaa  
 FT substituted with Ile"  
 FT Misc-difference 376 /label= Unknown  
 FT Misc-difference 394 /label= Unknown  
 FT Misc-difference 404 /label= Unknown  
 FT Misc-difference 413 /label= Unknown  
 FT  
 XX WO200161356-A1.  
 PN 23-AUG-2001.  
 XX  
 XX 16-FEB-2001; 2001WO-US005327.  
 XX  
 XX 16-FEB-2000; 2000US-00506079.  
 XX  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 XX Clinton G, Henner WD, Evans A;  
 XX  
 XX WPI; 2001-529934/58.  
 XX  
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the  
 XX treatment of hard tumors.  
 XX  
 XX Example 12; Page; 61pp; English.  
 XX  
 XX The invention relates to novel HER-2 (herstatin-2) antagonist  
 XX particularly a polypeptide that binds to the extracellular domain (ECD)  
 XX of HER-2 at a site that is different from the binding site of humanised  
 XX antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 XX invention is based upon the initial discovery of an alternative HER-2  
 XX mRNA transcript with 274 bp insert of intron 8. The translation product  
 XX of the alternative transcript is a truncated HER-2 protein designated  
 XX p68HER-2 which lacks the transmembrane and intracellular domains of  
 XX p68HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.  
 XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise  
 XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 XX nucleic acids encoding these are useful to treat, diagnose and identify  
 XX solid tumours. The present sequence is human p68HER-2 generic protein  
 XX containing ECDIIIA variant sequence. Note: The present sequence is not  
 XX shown in the specification but is derived from p68HER-2 generic sequence  
 XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)  
 XX  
 XX SQ Sequence 419 AA;  
 XX  
 XX Query Match 97.08; Score 128; DB 4; Length 419;  
 XX Best Local Similarity 95.8; Pred. No. 7.4e-11;  
 XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX  
 OY 1 PLNQEVTAEDGTORAEKSKPCA 24  
 DB 316 PLNQEVTAEDGTORAEKSKPCA 339  
 XX  
 XX RESULT 17  
 XX AAE09215  
 XX ID AAE09215 standard; protein; 419 AA.  
 XX  
 XX AC AAE09215;  
 XX  
 XX DT 15-NOV-2001 (first entry)

CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIa.  
 CC The ECDIIIa-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIa variant sequence. Note: The present sequence is not  
 CC shown in the specification but is derived from p68HER-2 generic sequence  
 CC (SEQ ID NO:13) shown in the sequence listing (AAE09183)  
 XX  
 SQ Sequence 419 AA;  
 Query Match 97.0%; Score 128; DB 4; Length 419;  
 Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 , : F:LNQEVTAEDGTQRAEKCSPCA 24  
 Db 316 PLHNQEVTAEDGTQRCCKSPCA 339  
 RESULT 18  
 AAE09181  
 ID AAE09216 standard; protein; 419 AA.  
 AC AAE09216;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 DE Human p68HER-2 generic protein variant 15.  
 XX  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIa; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the sequence shown in the  
 FT specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the sequence shown in the  
 FT specification"  
 FT Domain 341..419  
 FT /label= ECDIIIa variant  
 FT /note= "Extracellular domain IIIa"  
 FT Misc-difference 342  
 FT /label= Unknown  
 FT Misc-difference 345  
 FT /label= Unknown  
 FT Misc-difference 346  
 FT /label= Unknown  
 FT Misc-difference 356  
 FT /label= Unknown  
 FT Misc-difference 357  
 FT /label= Unknown  
 FT Misc-difference 358  
 FT /label= Unknown  
 FT Misc-difference 361  
 FT /label= Unknown  
 FT Misc-difference 371  
 FT /label= Unknown  
 FT Misc-difference 376  
 FT /label= Unknown  
 FT Misc-difference 394  
 FT /label= Unknown  
 FT Misc-difference 404  
 FT /label= Unknown  
 FT Misc-difference 413  
 FT /note= "p68HER-2 generic sequence (AAE09183) Xaa

FT substituted with Asn"  
 XX WO200161356-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 16-FEB-2001; 2001WO-US005327.  
 XX  
 PR 16-FEB-2000; 2000US-00506079.  
 XX  
 PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX  
 PI Clinton G, Henner WD, Evans A;  
 XX WPI; 2001-529934/58.  
 XX  
 PT New polypeptide, which binds to the extracellular domain of HER-2 for the  
 FT treatment of hard tumors.  
 XX  
 PS Example 12; Page; 61pp; English.  
 XX  
 CC The invention relates to novel HER-2 (herstatin-2) antagonist  
 CC particularly a polypeptide that binds to the extracellular domain (ECD)  
 CC of HER-2 at a site that is different from the binding site of humanised  
 CC antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present  
 CC invention is based upon the initial discovery of an alternative HER-2  
 CC mRNA transcript with 274 bp insert of intron 8. The translation product  
 CC of the alternative transcript is a truncated HER-2 protein designated  
 CC p68HER-2 which lacks the transmembrane and intracellular domains of  
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIa.  
 CC The ECDIIIa-containing polypeptides bind tightly to, and thus antagonise  
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the  
 CC nucleic acids encoding these are useful to treat, diagnose and identify  
 CC solid tumours. The present sequence is human p68HER-2 generic protein  
 CC containing ECDIIIa variant sequence. Note: The present sequence is not  
 CC shown in the specification but is derived from p68HER-2 generic sequence  
 CC (SEQ ID NO:13) shown in the sequence listing (AAE09183)  
 XX  
 SQ Sequence 419 AA;  
 Query Match 97.0%; Score 128; DB 4; Length 419;  
 Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24  
 Db 316 PLHNQEVTAEDGTQRCCKSPCA 339  
 RESULT 19  
 AAE09181  
 ID AAE09181 standard; protein; 419 AA.  
 XX  
 AC AAE09181;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 DE Human p68HER-2 generic sequence #1.  
 XX  
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;  
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;  
 KW ECDIIIa.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..340  
 FT /note= "Identical to N-terminal region of p185HER-2"  
 FT Misc-difference 124  
 FT /note= "Represented as Agn in the sequence shown in the  
 FT specification"  
 FT Misc-difference 125  
 FT /note= "Represented as Agn in the sequence shown in the





```

FT Misc-difference 413 /label= Unknown
FT FT /note= "Encoded by VAC"
XX
XX PN WO200161356-A1.
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US005327.
XX
XX PR 16-FEB-2000; 2000US-00506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX DR WPI; 2001-529934/58.
XX
XX DR N-PSDB; AADI5852.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX PS Claim 8; Page 57-58; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 protein containing
XX ECDIIIA generic sequence
XX
XX SQ Sequence 419 AA;

Query Match 97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24
DB 316 PLHNQEVTAEDGTQRAEKCKPCA 339

RESULT 21
AAE09200
ID AAE09200 standard; protein; 419 AA.
XX
XX AC AAE09200;
XX
XX DT 15-NOV-2001 (first entry)
XX
XX DE Human p68HER-2 generic protein variant (Arg357Cys).
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT Region 1..340
FT FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT FT /note= "Represented as Agn in the parent sequence shown
FT FT in the specification"
FT FT Misc-difference 125

```

```

FT FT /note= "Represented as Agn in the parent sequence shown
FT FT in the specification"
FT FT 341..419
FT FT /label= ECDIIIA variant
FT FT /note= "Extracellular domain IIIa variant"
FT FT Misc-difference 342 /label= Unknown
FT FT Misc-difference 345 /label= Unknown
FT FT Misc-difference 346 /label= Unknown
FT FT Misc-difference 356 /label= Unknown
FT FT Misc-difference 357 /label= Unknown
FT FT /note= "p68HER-2 generic sequence (AAE09181) Arg
FT FT substituted with Cys"
FT FT Misc-difference 358 /label= Unknown
FT FT Misc-difference 361 /label= Unknown
FT FT Misc-difference 376 /label= Unknown
FT FT Misc-difference 394 /label= Unknown
FT FT Misc-difference 404 /label= Unknown
FT FT Misc-difference 413 /label= Unknown
XX
XX WO200161356-A1.
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US005327.
XX
XX PR 16-FEB-2000; 2000US-00506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX DR WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX PS Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic sequence
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX SQ Sequence 419 AA;

Query Match 97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 7.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24

```

Db 316 PLHNOEVTADGTORCEKCKSKPCA 339

RESULT 22

AAE09204

ID AAE09204 standard; protein; 419 AA.

XX AC AAE09204;

XX DT 15-NOV-2001 (first entry)

XX DE Human p68HER-2 generic protein variant 2.

XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;

XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;

XX KW ECDIIIA; variant.

XX OS Homo sapiens.

XX FH Key

XX FH Region 1. .340

XX FH Location/Qualifiers

FT FT /note= "Identical to N-terminal region of p185HER-2"

FT FT Misc-difference 124

FT FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT FT Misc-difference 125

FT FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT FT Domain 341. .419

FT FT /label= ECDIIIA variant

FT FT /note= "Extracellular domain IIIA variant"

FT FT Misc-difference 342

FT FT /label= Unknown

FT FT Misc-difference 345

FT FT /label= Unknown

FT FT Misc-difference 346

FT FT /label= Unknown

FT FT Misc-difference 356

FT FT /label= Unknown

FT FT Misc-difference 358

FT FT /label= Unknown

FT FT Misc-difference 361

FT FT /label= Unknown

FT FT Misc-difference 376

FT FT /label= Unknown

FT FT Misc-difference 394

FT FT /label= Unknown

FT FT Misc-difference 404

FT FT /label= Unknown

FT FT Misc-difference 413

FT FT /label= Unknown

XX WO200161356-A1.

XX XX 23-AUG-2001.

XX XX 16-FEB-2001; 2001WO-US005327.

XX XX 16-FEB-2000; 2000US-00506079.

XX XX (UYOR-) UNIV OREGON HEALTH SCI.

XX XX Clinton G, Henner WD, Evans A;

XX XX WPI; 2001-529934/58.

XX XX New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors.

XX XX Example 11; Page; 61pp; English.

XX XX The invention relates to novel HER-2 (herstatin-2) antagonist

CC particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10<sup>8</sup>. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIIA variant sequence. Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181)

XX SQ Sequence 419 AA;

Query Match 97.0%; Score 128; DB 4; Length 419;

Best Local Similarity 95.8%; Pred. No. 7.4e-11;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNOEVTADGTORCEKCKSKPCA 24

Db 316 PLHNOEVTADGTORCEKCKSKPCA 339

|||||||

RESULT 23

AAE09209

ID AAE09209 standard; protein; 419 AA.

XX AC AAE09209;

XX DT 15-NOV-2001 (first entry)

XX DE Human p68HER-2 generic protein variant 7.

XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;

XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;

XX KW ECDIIIA; variant.

XX OS Homo sapiens.

XX FH Key

XX FH Region 1. .340

XX FH Location/Qualifiers

FT FT /note= "Identical to N-terminal region of p185HER-2"

FT FT Misc-difference 124

FT FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT FT Misc-difference 125

FT FT /note= "Represented as Agn in the parent sequence shown in the specification"

FT FT Domain 341. .419

FT FT /label= ECDIIIA variant

FT FT /note= "Extracellular domain IIIA variant"

FT FT Misc-difference 342

FT FT /label= Unknown

FT FT Misc-difference 345

FT FT /label= Unknown

FT FT Misc-difference 346

FT FT /label= Unknown

FT FT Misc-difference 356

FT FT /label= Unknown

FT FT Misc-difference 358

FT FT /label= Unknown

FT FT Misc-difference 361

FT FT /label= Unknown

FT FT Misc-difference 376

FT FT /label= Unknown

FT FT Misc-difference 394

FT FT /label= Unknown

FT FT Misc-difference 404

FT FT /label= Unknown

FT FT Misc-difference 413

FT FT /label= Unknown

XX WO200161356-A1.

XX XX 23-AUG-2001.

XX XX 16-FEB-2001; 2001WO-US005327.

XX XX 16-FEB-2000; 2000US-00506079.

XX XX (UYOR-) UNIV OREGON HEALTH SCI.

XX XX Clinton G, Henner WD, Evans A;

XX XX WPI; 2001-529934/58.

XX XX New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors.

XX XX Example 11; Page; 61pp; English.

XX XX The invention relates to novel HER-2 (herstatin-2) antagonist

```
FT FT /label= Unknown
FT FT Misc-difference 413
FT FT /label= Unknown
AX AX
PN PN WO200161356-A1.
PD PD 23-AUG-2001.
XX XX
XX XX 16-FEB-2001; 2001WO-US005327.
XX XX
XX XX 16-FEB-2000; 2000US-00506079.
XX XX
XX XX (UYOR-) UNIV OREGON HEALTH SCI.
XX XX
XX XX Clinton G, Henner WD, Evans A;
XX XX
XX XX WPI; 2001-529934/58.
XX XX
XX XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX XX treatment of hard tumors.
XX XX
XX XX Example 11; Page; 61pp; English.
XX XX
XX XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX XX of HER-2 at a site that is different from the binding site of humanised
XX XX antibody, Herceptin, at an affinity of at least 108. The present
XX XX invention is based upon the initial discovery of an alternative HER-2
XX XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX XX of the alternative transcript is a truncated HER-2 protein designated
XX XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX XX nucleic acids encoding these are useful to treat, diagnose and identify
XX XX solid tumours. The present sequence is human p68HER-2 generic protein
XX XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX XX shown in the specification but is derived from p68HER-2 generic sequence
XX XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX XX
XX XX Sequence 419 AA;
XX XX
XX XX Query Match 97.0%; Score 128; DB 4; Length 419;
XX XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX
OY 1 PLHNOEVTAEDGTQRAEKCCKPCA 24
DB 316 PLHNOEVTAEDGTQRCCKCKPCA 339
XX XX
RESULT 24
AAE20348
ID AAE20348 standard; protein; 419 AA.
XX XX
XX XX AAE20348;
XX XX
XX XX 18-JUN-2002 (first entry)
XX XX
XX XX Human truncated HER2 protein.
XX XX
XX XX Human; tumour; endothelial growth factor receptor; EGFR; cytostatic;
XX XX herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung;
XX XX colon; glial cell tumour; cell growth.
XX XX
XX XX Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX XX Misc-difference 342
XX XX /label= Thr, Ser
XX XX
XX XX Misc-difference 345
XX XX /label= Leu, Pro
XX XX Misc-difference 346
XX XX
```

---

```
FT FT /label= Pro, Leu
FT FT Misc-difference 356
FT FT /label= Leu, Gln
FT FT Misc-difference 358
FT FT /label= Met, Leu
FT FT Misc-difference 361
FT FT /label= Gly, Asp, Ala, Val
FT FT Misc-difference 376
FT FT /label= Leu, Ile
FT FT Misc-difference 394
FT FT /label= Pro, Arg
FT FT Misc-difference 404
FT FT /label= Pro, Leu
FT FT Misc-difference 413
FT FT /label= Asp, Asn
XX XX
XX XX WO200214470-A2.
XX XX
XX XX 21-FEB-2002.
XX XX
XX XX 14-AUG-2001; 2001WO-US025502.
XX XX
XX XX 14-AUG-2000; 2000US-00638834.
XX XX
XX XX (UYOR-) UNIV OREGON HEALTH SCI.
XX XX
XX XX Clinton GM;
XX XX
XX XX WPI; 2002-269185/31.
XX XX
XX XX Treating solid tumor characterized by expression of endothelial growth
XX XX factor receptor, involves administering recombinant herstatin that binds
XX XX to extracellular domain of the endothelial growth factor receptor.
XX XX
XX XX Claim 1; Page 78-80; 82pp; English.
XX XX
XX XX The present invention relates to a method for treating a solid tumour
XX XX characterised by endothelial growth factor receptor (EGFR) expression.
XX XX The method involves administering an agent that binds to an extracellular
XX XX domain (ECD) of EGFR. The invention also relates to a naturally occurring
XX XX inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
XX XX expression of herstatin with p185HER2 causes a striking reduction in cell
XX XX growth that corresponds with suppression of p185 autophosphorylation. The
XX XX method or a pharmaceutical composition is useful for treating a solid
XX XX tumour (selected from squamous cell carcinoma, lung carcinoma, colon
XX XX carcinoma and glial cell tumour) characterised by EGFR expression. The
XX XX present sequence is human truncated HER2 protein that lacks transmembrane
XX XX and intracellular domains
XX XX
XX XX Sequence 419 AA;
XX XX
XX XX Query Match 97.0%; Score 128; DB 5; Length 419;
XX XX Best Local Similarity 95.8%; Pred. No. 7.4e-11;
XX XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX
OY 1 PLHNOEVTAEDGTQRAEKCCKPCA 24
DB 316 PLHNOEVTAEDGTQRCCKCKPCA 339
XX XX
RESULT 25
ADP16709
ID ADP16709 standard; protein; 419 AA.
XX XX
XX XX ADP16709;
XX XX
XX XX 12-FEB-2004 (first entry)
XX XX
XX XX Human albumin fusion protein-related protein Segid1811.
XX XX
XX XX albumin fusion protein; albumin activity; human serum albumin;
XX XX serum osmotic pressure; shelf-life; stability; antidiabetic;
XX XX gene therapy; diabetes mellitus; human; gene; ds.
```

XX OS Homo sapiens.  
XX PN WO2003060071-A2.  
XX PD 24-JUL-2003.  
XX PF 23-DEC-2002; 2002WO-US040891.  
XX PR 21-DEC-2001; 2001US-0341811P.  
XX PR 24-JAN-2002; 2002US-0350358P.  
XX PR 28-JAN-2002; 2002US-0351360P.  
XX PR 26-FEB-2002; 2002US-0359370P.  
XX PR 28-FEB-2002; 2002US-0360000P.  
XX PR 27-MAR-2002; 2002US-0367500P.  
XX PR 08-APR-2002; 2002US-0370227P.  
XX PR 10-MAY-2002; 2002US-0378950P.  
XX PR 24-MAY-2002; 2002US-0382617P.  
XX PR 28-MAY-2002; 2002US-0383123P.  
XX PR 05-JUN-2002; 2002US-0385708P.  
XX PR 10-JUL-2002; 2002US-0394625P.  
XX PR 24-JUL-2002; 2002US-0398008P.  
XX PR 09-AUG-2002; 2002US-0402131P.  
XX PR 13-AUG-2002; 2002US-0402708P.  
XX PR 18-SEP-2002; 2002US-0411355P.  
XX PR 18-SEP-2002; 2002US-0411426P.  
XX PR 02-OCT-2002; 2002US-0414984P.  
XX PR 11-OCT-2002; 2002US-0417611P.  
XX PR 23-OCT-2002; 2002US-0420246P.  
XX PR 05-NOV-2002; 2002US-0423623P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.  
XX PA (PRIN-) PRINCIPIA PHARM CORP.  
XX PI Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;  
XX WPI; 2003-598517/56.  
XX DR N-PSDB; ADF16383.  
XX PT New albumin fusion protein, useful for preparing a composition for  
XX treating diabetes mellitus.  
XX Example 4; SEQ ID NO 1811; 24pp; English.  
XX This invention relates to a novel albumin fusion protein having albumin  
XX or biological activity. Human serum albumin is responsible for a  
XX significant proportion of the osmotic pressure of serum and also  
XX functions as a carrier of endogenous and exogenous ligands. The fusion of  
XX albumin to a therapeutic protein may increase shelf-life and stability of  
XX the therapeutic protein. The albumin fusion protein of the invention may  
XX allow production of compositions with antidiabetic activity whilst the  
XX nucleotide sequence which encodes it may be useful for gene therapy. The  
XX albumin fusion protein is useful for preparing a composition for treating  
XX diabetes mellitus. The present sequence is that of a therapeutic protein  
XX which was fused with human albumin to create a novel albumin fusion  
XX protein of the invention. Note: The sequence data for this patent did not  
XX form part of the printed specification, but was obtained in electronic  
XX format directly from WIPO at ftp.wipo.int/pub/publishedpct\_sequences

XX SQ Sequence 419 AA;  
Query Match 97.0%; Score 128; DB 7; Length 419;  
Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 26  
ADY28119

ID ADY28119 standard; protein; 419 AA.  
XX AC ADY28119;  
XX DT 05-MAY-2005 (first entry)  
XX DE Receptor tyrosine kinase protein - SEQ ID 9.  
XX KW receptor tyrosine kinase; angiogenesis; cell proliferation;  
KW cell migration; cell growth; metastasis; cancer; cytostatic;  
KW rheumatoid arthritis; antiarthritic; antirheumatic; multiple sclerosis;  
KW neuroprotective; inflammation; antiinflammatory; uveitis;  
KW ophthalmological; ocular inflammation; atherosclerosis;  
KW antiarteriosclerotic; hemangioma; diabetes mellitus;  
KW inflammatory bowel disease; gastrointestinal-gen.; Crohns disease;  
KW psoriasis; antipsoriatic; Alzheimers disease; neuroprotective; nootropic;  
KW systemic lupus erythematosus; restenosis; vasotropic; asthma;  
KW antiasthmatic; receptor tyrosine kinase.  
XX OS Unidentified.  
XX PN WO2005016966-A2.  
XX PD 24-FEB-2005.  
XX PF 14-MAY-2004; 2004WO-US015056.  
XX PR 16-MAY-2003; 2003US-0471141P.  
XX PA (RECE-) RECEPTOR BIOLOGIX INC.  
XX PA (JINP/) JIN P.  
XX PI Jin P, Shepard HM, Clinton GM, Lackey DB;  
XX WPI; 2005-182332/19.  
XX PT New receptor tyrosine kinase isoforms (such as TEK (TIE-2)), useful for  
XX treating a disease or condition, e.g. cancers, or inflammatory,  
XX infectious, immune, or neurodegenerative diseases.  
XX Example 1; SEQ ID NO 9; 105pp; English.  
XX The invention comprises the amino acid sequence of a number of receptor  
XX tyrosine kinase proteins which are useful for identifying drugs that  
XX modulate the receptor and for inhibiting angiogenesis, cell  
XX proliferation, cell migration, tumor cell growth or tumor cell  
XX metastasis. The receptor tyrosine kinase proteins of the invention are  
XX useful for treating cancers, rheumatoid arthritis, multiple sclerosis and  
XX posterior intraocular inflammation, uveitic disorders, ocular surface  
XX inflammatory disorders, neovascular disease, proliferative  
XX vitreoretinopathy, atherosclerosis, hemangioma, diabetes mellitus,  
XX inflammatory bowel disease, Crohn's disease, psoriasis, Alzheimer's  
XX disease, lupus, vascular stenosis, restenosis, inflammatory joint  
XX disease, urinary obstructive syndromes, and asthma. The present amino  
XX acid sequence represents a receptor tyrosine kinase protein of the  
XX invention. NOTE: The present sequence is not shown in the specification,  
XX but has been retrieved from the WIPO website.  
XX SQ Sequence 419 AA;  
Query Match 97.0%; Score 128; DB 9; Length 419;  
Best Local Similarity 95.8%; Pred. No. 7.4e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 27  
ADY28119

AC	AD247861;	XX	Key	Location/Qualifiers
XX		XX	Misc-difference 124	/note= "represented as Agn"
DT	30-JUN-2005 (first entry)	FT	Misc-difference 125	/note= "represented as Agn"
XX		FT	Misc-difference 125	/note= "represented as Agn"
DE	Wild-type human ErbB-2 protein.	FT	Misc-difference 343	/note= "Preferably Ser"
XX		FT	Misc-difference 346	/note= "Preferably Pro"
KW	epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis; prognosis; cancer; cytosstatic.	FT	Misc-difference 347	/note= "Preferably Leu"
XX		FT	Misc-difference 357	/note= "Preferably Gln"
OS	Homo sapiens.	FT	Misc-difference 359	/note= "Preferably Leu"
XX		FT	Misc-difference 362	/note= "changes from glycine"
PN	WO2005033133-A2.	FT	Misc-difference 377	/note= "Preferably Ile"
XX		FT	Misc-difference 395	/note= "Preferably Arg"
PD	14-APR-2005.	FT	Misc-difference 405	/note= "Preferably Leu"
XX		FT	Misc-difference 414	/note= "Preferably Asn"
PF	04-OCT-2004; 2004WO-US030903.	XX	WO20044403-A1.	
XX		PN		
PR	03-OCT-2003; 2003US-0507953P.	XX	03-AUG-2000.	
PR	10-NOV-2003; 2003US-0518321P.	XX	20-JAN-2000; 2000WO-US001484.	
PR	09-AUG-2004; 2004US-0599583P.	PR	20-JAN-1999; 99US-00234208.	
PR	07-SEP-2004; 2004US-0607326P.	XX	(UYOR-) UNIV OREGON HEALTH SCI.	
XX	(COMP-) COMPUGEN LTD.	XX	Doherty JK, Clinton GM, Adelman JP;	
PA	(COMP-) COMPUGEN INC.	XX	WPI; 2000-499287/44.	
XX	Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;	XX	Using polypeptides and antibodies that bind to the extracellular domain of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the breast, lung, ovaries and colon.	
PI	Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;	PS	Claim 8; Page 39-40; 46pp; English.	
XX		XX	HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The extracellular domain of p185-HER-2 is proteolytically shed from breast carcinoma cells in culture and is found in serum of some cancer patients and may be a serum marker of metastatic breast cancer. An alternative HER-2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been identified. The retained intron is in-frame and encodes a 79 amino acid extension designated ECDIIa (the present sequence), which is inserted at residue 340 of p185-HER-2. The alternative mRNA predicts a truncated HER-2 protein (approximately 68 kDa) that lacks the transmembrane and intracellular domains (see AAY97240). p68HER-2 specifically binds to p185-HER-2 without activating HER-2. It could therefore block dimerization of p185-HER-2. The p68HER-2 polypeptide binds to a site on the ECD of HER-2 that is different from the site of binding for Herceptin (RTM) (a marketed humanized monoclonal antibody that is used for the treatment of cancer and binds to the ECD of HER-2). The methods, compositions, polypeptides and antibodies are used to treat solid tumours such as breast cancer, small cell lung carcinoma, ovarian cancer and/or colon cancer, especially where over-expression of HER-2 is indicated	
DR	WPI; 2005-285403/29.	XX	Sequence 420 AA;	
XX		XX	Query Match 97.0%; Score 128; DB 3; Length 420;	
PT	New isolated polynucleotide encoding at least an active portion of an ErbB-2 polypeptide, useful for diagnosing or treating diseases with aberrant expression or activity of the ErbB-2 polypeptides, such as cancer.	XX	Best Local Similarity 95.8%; Pred. No. 7.4e-11;	
PT		XX	Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
PS	Disclosure; Fig 4; 246pp; English.	QY	1 PLHNQEVTAEDGTQRAEKCSPCA 24	
XX		DB	316 PLHNQEVTAEDGTQRAEKCSPCA 339	
CC	The invention relates to the isolation of polynucleotide sequences encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as HER-2 or neu) polypeptides. Also described are methods, compositions, and kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and treatment of ErbB-2 related cancer. The sequences, methods, and compositions of the present invention are useful for diagnosing, prognosing, and/or treating diseases or conditions associated with aberrant expression or activity of the ErbB-2 polypeptides, such as cancer. This sequence represents the wild-type human ErbB-2 protein.	RESULT 28		
XX		AAY97240		
CC		ID AAY97240 standard; protein; 420 AA.		
CC		AC AAY97240;		
DT	04-DEC-2000 (first entry)	XX		
XX		DE Truncated HER-2, p68-HER-2.		
XX		KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion; extracellular domain IITa; antagonist; intron 8; C-terminal extension;		
KW	truncated HER-2; p68; dimerization inhibitor; cytosstatic.	XX		
XX		OS Homo sapiens.		

```
Db      316 PLHNQEVTAEDGTQCEKCKPCA 339
|||||
RESULT 29
AAE36793
ID AAE36793 standard; protein; 509 AA.
AC AAE36793;
XX
XX 07-AUG-2003 (first entry)
XX
XX Human ErbB-2 protein.
XX
XX Epidermal growth factor receptor; EGFR; therapy; psoriasis; carcinoma;
XX cancer; rhabdomyosarcoma; mesothelioma; melanoma; glioblastoma; human;
XX receptor; ErbB-2.
XX
XX Homo sapiens.
XX
XX WO2003014159-A1.
XX
XX 20-FEB-2003.
XX
XX 05 AUG 2002; 2002WO-AU001042.
XX
XX 03-AUG-2001; 2001AU-00006927.
XX
XX 03-AUG-2001; 2001AU-00006928.
XX
XX 01-NOV-2001; 2001US-0335333P.
XX
XX 01-NOV-2001; 2001US-0336560P.
XX
XX 31-MAY-2002; 2002AU-00002731.
XX
XX 11-JUN-2002; 2002US-0388171P.
XX
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX (BIOM-) BIOMOLECULAR RES INST LTD.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Adams TE, Burgess AW, Elleman TC, Garrett TPJ, Jorissen RN;
XX Lou M, Lovrecz GO, Mckern NM, Nice EC, Ward CW;
XX WPI; 2003-268181/26.
XX
XX Selecting or designing compounds that interact with or inhibit formation
XX of active dimers of the EGF receptor family, and useful for the
XX prevention and treatment of disorders, such as psoriasis and cancer of
XX the breast, brain or colon.
XX
XX Disclosure; Fig 1; 354pp; English.
XX
XX The invention relates to a method of selecting or designing a compound
XX that interacts with or inhibits the formation of active dimers of a
XX receptor of the epidermal growth factor receptor (EGFR) family. The
XX methods and compositions of the invention are useful for the prevention
XX and treatment of disorders associated with signalling by a molecule of
XX the EGFR family such as psoriasis and cancer of the pancreas, breast,
XX brain, colon, prostate, ovary, cervix, lung, head and neck, melanoma,
XX rhabdomyosarcoma, mesothelioma, squamous carcinomas of the skin and
XX glioblastomas. The present sequence is human ErbB-2 protein. This
XX sequence is used to illustrate the method of the invention
XX
XX Sequence 509 AA;
Query Match 97.0%; Score 128; DB 6; Length 509;
Best Local Similarity 95.8%; Pred. No. 9.1e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24
|||||
Db 295 PLHNQEVTAEDGTQCEKCKPCA 318
|||||
RESULT 30
AAE36793
ID AAE36793 standard; protein; 509 AA.
AC AAE36793;
XX
XX 07-AUG-2003 (first entry)
XX
XX Human ErbB-2 protein.
XX
XX Epidermal growth factor receptor; EGFR; therapy; psoriasis; carcinoma;
XX cancer; rhabdomyosarcoma; mesothelioma; melanoma; glioblastoma; human;
XX receptor; ErbB-2.
XX
XX Homo sapiens.
XX
XX WO2003014159-A1.
XX
XX 20-FEB-2003.
XX
XX 05 AUG 2002; 2002WO-AU001042.
XX
XX 03-AUG-2001; 2001AU-00006927.
XX
XX 03-AUG-2001; 2001AU-00006928.
XX
XX 01-NOV-2001; 2001US-0335333P.
XX
XX 01-NOV-2001; 2001US-0336560P.
XX
XX 31-MAY-2002; 2002AU-00002731.
XX
XX 11-JUN-2002; 2002US-0388171P.
XX
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX (BIOM-) BIOMOLECULAR RES INST LTD.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Adams TE, Burgess AW, Elleman TC, Garrett TPJ, Jorissen RN;
XX Lou M, Lovrecz GO, Mckern NM, Nice EC, Ward CW;
XX WPI; 2003-268181/26.
XX
XX Selecting or designing compounds that interact with or inhibit formation
XX of active dimers of the EGF receptor family, and useful for the
XX prevention and treatment of disorders, such as psoriasis and cancer of
XX the breast, brain or colon.
XX
XX Disclosure; Fig 1; 354pp; English.
XX
XX The invention relates to a method of selecting or designing a compound
XX that interacts with or inhibits the formation of active dimers of a
XX receptor of the epidermal growth factor receptor (EGFR) family. The
XX methods and compositions of the invention are useful for the prevention
XX and treatment of disorders associated with signalling by a molecule of
XX the EGFR family such as psoriasis and cancer of the pancreas, breast,
XX brain, colon, prostate, ovary, cervix, lung, head and neck, melanoma,
XX rhabdomyosarcoma, mesothelioma, squamous carcinomas of the skin and
XX glioblastomas. The present sequence is human ErbB-2 protein. This
XX sequence is used to illustrate the method of the invention
XX
XX Sequence 509 AA;
Query Match 97.0%; Score 128; DB 6; Length 509;
Best Local Similarity 95.8%; Pred. No. 9.1e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24
|||||
Db 295 PLHNQEVTAEDGTQCEKCKPCA 318
|||||
RESULT 31
AAE36793
ID AAE36793 standard; protein; 509 AA.
AC AAE36793;
XX
XX 01-JUL-2004 (first entry)
XX
XX Human ErbB2 polypeptide ectodomain fragment.
XX
XX ErbB2; EGF receptor; ErbB1; ErbB5; ErbB4; cytostatic; human;
XX protein co-ordinate data.
XX
XX Homo sapiens.
XX
XX WO2004031232-A1.
XX
XX 15-APR-2004.
XX
XX 06-OCT-2003; 2003WO-AU001310.
XX
XX 04-OCT-2002; 2002AU-00951853.
XX
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX (LUDW-) LUDWIG INST CANCER RES.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
XX Garrett TPJ, Elleman TC, Adams TE, Burgess AW, Jorissen RN;
XX Lou M, Lovrecz GO, Mckern NM, Ward CW;
XX WPI; 2004-347960/32.
XX
XX Identifying and designing compounds interacting with ErbB2, used in
XX diagnosing diseases, involves assessing stereochemical complementarity
XX between three-dimensional structure of amino acids 1-509 of ErbB2
XX polypeptide and candidate compound.
XX
XX Disclosure; SEQ ID NO 1; 148pp; English.
XX
XX There invention relates to a method for identifying a potential modulator
XX compound for ErbB2 which involves assessing the stereochemical
XX complementarity of a three-dimensional structure of amino acids 1-509 of
XX ErbB2 polypeptide or their subsets and a region of the three-dimensional
XX structure of a candidate compound and selecting a compound on the basis
XX of the stereochemical complementarity. Identifying a potential modulator
XX compound for ErbB2 or the method of modulating ErbB2 further comprises
XX synthesizing or obtaining a candidate compound assessed as possessing
XX stereochemical complementarity with a topographical region of the three-
XX dimensional structure, and determining the ability of the candidate
XX compound to interact with and/or modulate the activity of ErbB2. The
XX subset of amino acids is selected from at least one of the CR1 domain,
XX the potential CR1 loop docking site between the L1, CR1 and L2 domains,
XX the CR1-L2 hinge region, the regions of the L1 and L2 domains that
XX contact each other in a closed conformation, and defines at least a part
XX of the heterodimerization surface with another member of the EGF receptor
XX family that is selected from the group consisting of ErbB1 (EGF
XX receptor), ErbB5 and ErbB4. The method is useful for screening for and
XX designing compounds that interact with ErbB2 or their variants, for
XX treating diseases such as cancer, melanoma, rhabdomyosarcoma,
XX mesothelioma and glioblastoma. The present sequence represents the
XX ectodomain of human ErbB2 protein.
XX
XX Sequence 509 AA;
Query Match 97.0%; Score 128; DB 8; Length 509;
Best Local Similarity 95.8%; Pred. No. 9.1e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24
|||||
Db 295 PLHNQEVTAEDGTQCEKCKPCA 318
|||||
RESULT 31
```

ADT50878  
ID ADT50878 standard; protein; 575 AA.  
XX  
AC ADT50878;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Cancer related protein sequence #41.  
XX  
DE cytostatic; gene therapy; vaccine; diagnosis; breast; colon; lung;  
KW ovarian; prostate; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO2004092338-A2.  
XX  
PD 28-OCT-2004.  
XX  
PF 12-APR-2004; 2004WO-US011104.  
XX  
PR 11-APR-2003; 2003US-0462399P.  
PR 01-JUL-2003; 2003US-0484333P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Macina RA, Turner LR, Sun Y, Tam A;  
XX  
DR WPI; 2004-766851/75.  
XX  
PT New cancer specific nucleic acid (CaSNA) molecules, useful for  
PT diagnosing, monitoring the presence of, or treating a patient with  
PT breast, colon, lung, ovarian, or prostate cancer.  
XX  
PS Claim 12; SEQ ID NO 182; 891pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I)  
CC selectively hybridizing to, or comprising at least 95% sequence identity  
CC to, any of the 362 nucleotide sequences fully defined in the  
CC specification. The nucleic acid molecules and polypeptides are useful for  
CC diagnosing, monitoring the presence of, or treating a patient with  
CC breast, colon, lung, ovarian, or prostate cancer. This sequence  
CC corresponds to a protein of the invention.  
XX  
SQ Sequence 575 AA;  
Query Match 97.0%; Score 128; DB 8; Length 575;  
Best Local Similarity 95.8%; Pred. No. 1e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
DB 316 PLHNQEVTAEDGTQRAEKCCKPCA 339  
RESULT 32  
ADY28117  
ID ADY28117 standard; protein; 575 AA.  
XX  
AC ADY28117;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE Receptor tyrosine kinase protein - SEQ ID 7.  
XX  
KW receptor tyrosine kinase; angiogenesis; cell proliferation;  
KW cell migration; cell growth; metastasis; cancer; cytosstatic;  
KW rheumatoid arthritis; antiarthritic; antirheumatic; multiple sclerosis;  
KW neuroprotective; inflammation; antiinflammatory; uveitis;  
KW ophthalmological; ocular inflammation; atherosclerosis;  
KW antiarteriosclerotic; hemangioma; diabetes mellitus;  
KW inflammatory bowel disease; gastrointestinal-gen.; Crohns disease;  
KW psoriasis; antipsoriatic; Alzheimers disease; neuroprotective; nootropic;  
KW systemic lupus erythematosus; restenosis; vasotropic; asthma;

KW antiasthmatic; receptor tyrosine kinase.  
XX  
OS Unidentified.  
XX  
PN WO2005016966-A2.  
XX  
PD 24-FEB-2005.  
XX  
PF 14-MAY-2004; 2004WO-US015056.  
XX  
PR 16-MAY-2003; 2003US-0471141P.  
XX  
PA (RECE-) RECEPTOR BIOLOGIX INC.  
XX (JINP/) JIN P.  
XX  
PI Jin P, Shepard HM, Clinton GM, Lackey DB;  
XX  
DR WPI; 2005-182332/19.  
XX  
PT New receptor tyrosine kinase isoforms (such as TEK (TIE-2)), useful for  
PT treating a disease or condition, e.g. cancers, or inflammatory,  
PT infectious, immune, or neurodegenerative diseases.  
XX  
PS Claim 2; SEQ ID NO 7; 105pp; English.  
XX  
CC The invention comprises the amino acid sequence of a number of receptor  
CC tyrosine kinase proteins which are useful for identifying drugs that  
CC modulate the receptor and for inhibiting angiogenesis, cell  
CC proliferation, cell migration, tumor cell growth or tumor cell  
CC metastasis. The receptor tyrosine kinase proteins of the invention are  
CC useful for treating cancers, rheumatoid arthritis, multiple sclerosis and  
CC posterior intraocular inflammation, uveitic disorders, ocular surface  
CC inflammatory disorders, neovascular disease, proliferative  
CC vitreoretinopathy, atherosclerosis, hemangioma, diabetes mellitus,  
CC inflammatory bowel disease, Crohn's disease, psoriasis, Alzheimer's  
CC disease, lupus, vascular stenosis, restenosis, inflammatory joint  
CC disease, urinary obstructive syndromes, and asthma. The present amino  
CC acid sequence represents a receptor tyrosine kinase protein of the  
CC invention. NOTE: The present sequence is not shown in the specification,  
CC but has been retrieved from the WIPO website.  
XX  
SQ Sequence 575 AA;  
Query Match 97.0%; Score 128; DB 9; Length 575;  
Best Local Similarity 95.8%; Pred. No. 1e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
DB 316 PLHNQEVTAEDGTQRAEKCCKPCA 339  
RESULT 33  
ADZ47797  
ID ADZ47797 standard; protein; 575 AA.  
XX  
AC ADZ47797;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Amino acid sequence for human ErbB-2 variant II.  
XX  
KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH Active-site 505..575  
FT /note= "Given as SEQ ID No:6 in the specification"  
XX  
PN WO2005033133-A2.  
XX



PD 14-APR-2005.  
XX  
PF 04-OCT-2004; 2004WO-US030903.  
XX  
PR 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX  
PA (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
DR WPI; 2005-285403/29.  
DR N-PSDB; ADZ47796.  
XX  
XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
PT cancer.  
XX  
XX Claim 19; SEQ ID NO 4; 246pp; English.  
XX  
CC The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
CC cancer. This sequence represents human ErbB-2 variant II.  
XX  
XX SQ Sequence 575 AA;  
Query Match 97.0%; Score 128; DB 9; Length 575;  
Best Local Similarity 95.8%; Pred. No. 1e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339  
RESULT 34  
ADZ47809  
ID ADZ47809 standard; protein; 614 AA.  
XX  
AC ADZ47809;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Amino acid sequence for human ErbB-2 variant V.  
XX  
KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200503133-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 04-OCT-2004; 2004WO-US030903.  
XX  
PR 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX  
PA (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.

PA (COMP-) COMPUGEN INC.  
XX  
PI Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
DR WPI; 2005-285403/29.  
DR N-PSDB; ADZ47808.  
XX  
XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
PT cancer.  
XX  
XX Claim 41; SEQ ID NO 16; 246pp; English.  
XX  
CC The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
CC cancer. This sequence represents human ErbB-2 variant V.  
XX  
XX SQ Sequence 614 AA;  
Query Match 97.0%; Score 128; DB 9; Length 614;  
Best Local Similarity 95.8%; Pred. No. 1e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339  
RESULT 35  
AAR08222  
ID AAR08222 standard; protein; 624 AA.  
XX  
AC AAR08222;  
XX  
DT 06-MAR-1991 (first entry)  
XX  
DE Extracellular portion of the human epidermal growth factor receptor 2.  
XX  
KW Human epidermal growth factor receptor 2; HER2; vaccine; cancer.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 22..32  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 74..84  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 113..130  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 187..197  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 305..315  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 327..353  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 375..385  
FT /label= epitope  
FT /note= "potential T-cell epitope"  
FT Region 398..408

FT /label= epitope  
 FT /note= "potential T-cell epitope"  
 FT 433..443  
 FT /label= epitope  
 FT /note= "potential T-cell epitope"  
 FT 519..529  
 FT /label= epitope  
 FT /note= "potential T-cell epitope"  
 XX  
 XX WO9014357-A.  
 XX  
 XX PN  
 XX PD 29-NOV-1990.  
 XX  
 XX PF 19-MAY-1989; 89US-00354319.  
 XX  
 XX PR 19-MAY-1989; 89US-00354319.  
 XX  
 XX PA (GETH ) GENENTECH INC.  
 XX  
 XX PI Hudziak RM, Shepard HM, Ullrich A;  
 XX  
 XX WPI; 1990-375946/50.  
 XX  
 XX DR N-PSDB; AAQ06828.  
 XX  
 XX HER2 extracellular domain used as vaccine - comprises sequence of at  
 PT least 9 aminoacid(s) prep. using expression vector of DNA isolated from  
 PT human epidermal growth factor receptor.  
 XX  
 XX PS Disclosure; Fig 13; 49pp; English.  
 XX  
 XX CC This claimed sequence is free of transmembrane or intracellular portions  
 CC of the HER2 mol. The protein is antigenic in animals. Fragment comprising  
 CC at least 9 amino acids open several possibilities for further research  
 CC and a broad range of potential therapeutic applications, e.g. for the  
 CC treatment of mammary gland adenocarcinoma and other cancers  
 XX  
 XX SQ Sequence 624 AA;  
 Query Match 97.0%; Score 128; DB 2; Length 624;  
 Best Local Similarity 95.8%; Pred. No. 1.1e 10;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24  
 |||||  
 DB 295 PLHNQEVTAEDGTQRAEKCKPKCA 318  
 |||||  
 RESULT 36  
 ADY28116  
 ID ADY28116 standard; protein; 633 AA.  
 XX  
 XX AC ADY28116;  
 XX  
 XX DT 05-MAY-2005 (first entry)  
 XX  
 XX DE Receptor tyrosine kinase protein - SEQ ID 6.  
 XX  
 XX KW receptor tyrosine kinase; angiogenesis; cell proliferation;  
 KW cell migration; cell growth; metastasis; cancer; cytostatic;  
 KW rheumatoid arthritis; anarthritic; antineumatic; multiple sclerosis;  
 KW neuroprotective; inflammation; antinflammatory; uveitis;  
 KW ophthalmological; ocular inflammation; atherosclerosis;  
 KW antiarteriosclerotic; hemangioma; diabetes mellitus;  
 KW inflammatory bowel disease; gastrointestinal-gen.; Crohns disease;  
 KW psoriasis; antipsoriatic; Alzheimers disease; neuroprotective; nootropic;  
 KW systemic lupus erythematosus; restenosis; vasotropic; asthma;  
 KW antiasthmatic; receptor tyrosine kinase.  
 XX  
 XX OS Unidentified.  
 XX  
 XX PN WO2005016966-A2.  
 XX  
 XX PD 24-FEB-2005.

XX  
 XX PF 14-MAY-2004; 2004WO-US015056.  
 XX  
 XX PR 16-MAY-2003; 2003US-0471141P.  
 XX  
 XX (RECE-) RECEPTOR BIOLOGIX INC.  
 XX (JINP/) JIN P.  
 XX  
 XX PI Jin P, Shepard HM, Clinton GM, Lackey DB;  
 XX  
 XX WPI; 2005-182332/19.  
 XX  
 XX PT New receptor tyrosine kinase isoforms (such as TEK (TIE-2)), useful for  
 PT treating a disease or condition, e.g. cancers, or inflammatory,  
 PT infectious, immune, or neurodegenerative diseases.  
 XX  
 XX PS Claim 6; SEQ ID NO 6; 105pp; English.  
 XX  
 XX CC The invention comprises the amino acid sequence of a number of receptor  
 CC tyrosine kinase proteins which are useful for identifying drugs that  
 CC modulate the receptor and for inhibiting angiogenesis, cell  
 CC proliferation, cell migration, tumor cell growth or tumor cell  
 CC metastasis. The receptor tyrosine kinase proteins of the invention are  
 CC useful for treating cancers, rheumatoid arthritis, multiple sclerosis and  
 CC posterior intraocular inflammation, uveitic disorders, ocular surface  
 CC inflammatory disorders, neovascular disease, proliferative  
 CC vitreoretinopathy, atherosclerosis, hemangioma, diabetes mellitus,  
 CC inflammatory bowel disease, Crohn's disease, psoriasis, Alzheimer's  
 CC disease, lupus, vascular stenosis, restenosis, inflammatory joint  
 CC disease, urinary obstructive syndromes, and asthma. The present amino  
 CC acid sequence represents a receptor tyrosine kinase protein of the  
 CC invention. NOTE: The present sequence is not shown in the specification,  
 CC but has been retrieved from the WIPO website.  
 XX  
 XX SQ Sequence 633 AA;  
 Query Match 97.0%; Score 128; DB 9; Length 633;  
 Best Local Similarity 95.8%; Pred. No. 1.2e-10;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24  
 |||||  
 DB 316 PLHNQEVTAEDGTQRAEKCKPKCA 339  
 |||||  
 RESULT 37  
 ADZ47862  
 ID ADZ47862 standard; protein; 633 AA.  
 XX  
 XX AC ADZ47862;  
 XX  
 XX DT 30-JUN-2005 (first entry)  
 XX  
 XX DE Human 100 KDa ErbB-2 variant polypeptide.  
 XX  
 XX KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
 KW prognosis; cancer; cytostatic.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO2005033133-A2.  
 XX  
 XX PD 14-APR-2005.  
 XX  
 XX PF 04-OCT-2004; 2004WO-US030903.  
 XX  
 XX PR 03-OCT-2003; 2003US-0507953P.  
 XX 10-NOV-2003; 2003US-0518321P.  
 XX PR 09-AUG-2004; 2004US-0599583P.  
 XX 07-SEP-2004; 2004US-0607326P.  
 XX  
 XX PA (COMP-) COMPUGEN LTD.  
 XX (COMP-) COMPUGEN INC.

XX Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
 PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
 XX WPI; 2005-285403/29.  
 DR  
 XX New isolated polynucleotide encoding at least an active portion of an  
 PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
 PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
 PT cancer.  
 XX  
 XX Disclosure; Fig 4; 246pp; English.  
 PS  
 XX The invention relates to the isolation of polynucleotide sequences  
 CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
 CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
 CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
 CC treatment of ErbB-2 related cancer. The sequences, methods, and  
 CC compositions of the present invention are useful for diagnosing,  
 CC prognosing, and/or treating diseases or conditions associated with  
 CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
 CC cancer. This sequence represents human 100 KDa ErbB-2 variant  
 CC polypeptide.  
 XX  
 XX Sequence 633 AA;  
 SQ

Query Match 97.0%; Score 128; DB 9; Length 633;  
 Best Local Similarity 95.8%; Pred. No. 1.2e-10;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
 |||||  
 DB 316 PLHNQEVTAEDGTQRCCKCKSPCA 339

RESULT 38  
 AAB60408  
 ID AAB60408 standard; protein; 645 AA.  
 XX  
 AC AAB60408;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE Human ErbB2 oncoprotein, SEQ ID NO:13.  
 XX  
 KW Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL;  
 KW light chain variable region; cancer; cytostatic; EGFR-expressing cancer;  
 KW epidermal growth factor receptor; colon cancer; rectal cancer; tumour;  
 KW colorectal cancer; non-small cell lung cancer; metastatic breast cancer;  
 KW affinity purification.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100245-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-US017366.  
 XX  
 PR 25-JUN-1999; 99US-0141316P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Adams CW, Presta LG, Sliwkowsky M;  
 PI  
 XX WPI; 2001-080862/09.  
 DR  
 XX Treating cancer in a human, where the cancer expresses epidermal growth  
 PT factor receptor (EGFR), comprises administering an antibody which binds  
 PT ErbB2.  
 XX  
 XX Example 1; Fig 1A; 89pp; English.  
 PS  
 XX

CC The invention relates to a method for treating cancer in a human patient,  
 CC wherein the cancer expresses epidermal growth factor receptor (EGFR),  
 CC comprising administering an antibody which binds ErbB2 (HER2; AAB60408).  
 CC In particular, the anti-ErbB2 antibody is the murine monoclonal antibody  
 CC 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398,  
 CC AAB60399). The invention also encompasses an isolated nucleic acid  
 CC encoding a humanised ErbB2-binding antibody; vectors and host cells  
 CC comprising such nucleic acids; the recombinant production of a humanised  
 CC ErbB2-binding antibody; and an immunoconjugate comprising a humanised  
 CC ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies  
 CC act by antagonising ErbB receptors, and as inhibitors of transforming  
 CC growth factor alpha (TGF-alpha)-activated mitogen activated protein  
 CC kinase (MAPK). The method of the invention is used for treating cancer,  
 CC especially colon cancer, rectal cancer, colorectal cancer, lung cancer,  
 CC (especially non-small cell lung cancer), or breast cancer (especially  
 CC metastatic breast cancer). The antibodies may also have non-therapeutic  
 CC uses e.g., as affinity purification agents. Using an antibody which binds  
 CC to ErbB2 to treat cancer is preferable to the use of EGFR-targeted  
 CC drugs, as EGFR is also highly expressed in other tissues such as the  
 CC liver and skin, where the active drug will also bind, with skin toxicity  
 CC having been observed for EGFR-targeted drugs. Antibodies which bind  
 CC ErbB2 are anticipated to have a better safety profile than such drugs.  
 CC The present sequence represents human ErbB2  
 XX  
 XX Sequence 645 AA;  
 SQ

Query Match 97.0%; Score 128; DB 4; Length 645;  
 Best Local Similarity 95.8%; Pred. No. 1.2e-10;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
 |||||  
 DB 316 PLHNQEVTAEDGTQRCCKCKSPCA 339

RESULT 39  
 AAB61593  
 ID AAB61593 standard; protein; 645 AA.  
 XX  
 AC AAB61593;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Human ErbB2 extracellular domain.  
 XX  
 KW Human; ErbB2; cytostatic; prostate cancer; receptor tyrosine kinase;  
 KW antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100238-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-US017423.  
 XX  
 PR 25-JUN-1999; 99US-0141315P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
 XX  
 XX Agus DB, Scher HI, Sliwkowski MX;  
 PI  
 XX WPI; 2001-159131/16.  
 DR  
 XX Treating prostate cancer in a human comprises administering an antibody  
 PT which binds ErbB2 and blocks ligand activation of an ErbB receptor.  
 PT  
 XX Disclosure; Fig 1; 93pp; English.  
 PS  
 XX The ErbB family of receptor tyrosine kinases are important mediators of  
 CC cell growth, differentiation and survival. The receptor family includes  
 CC four distinct members including Epidermal Growth Factor Receptor (EGFR) or

CC ErbB1), HER2 (ErbB2 or p185<sup>neu</sup>), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).  
CC The present invention relates to a method for treating prostate cancer.  
CC The method comprises administering an antibody which binds ErbB2 and  
CC blocks ligand activation of an ErbB receptor. Preferably, the antibody  
CC blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks TGF-  
CC alpha activation of mitogen-activated protein kinase (MAPK). The present  
CC sequence is the extracellular domain of human ErbB2

XX SQ Sequence 645 AA;

Query Match 97.0%; Score 128; DB 4; Length 645;

Best Local Similarity 95.8%; Pred. No. 1.2e-10; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24

|||||

316 PLHNQEVTAEDGTQRAEKCKSPCA 339

RESULT 40

ABG70753

ID ABG70753 standard; protein; 645 AA.

XX AC ABG70753;

XX DT 28-NOV-2002 (first entry)

XX DE Human HER2 receptor extracellular domain.

XX KW Human; HER2; analyte; interfering substance; serum; HERCEPTIN; therapy;

XX KW anti-HER2; antibody; plasma; HER2; ErbB2; ErbB2 receptor;

XX KW extracellular domain; ECD; epitope; cancer.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide

1..21

/label= Signal\_peptide

XX US2002090662-A1.

XX PD 11-JUL-2002.

XX PF 01-AUG-2001; 2001US-00921161.

XX PR 15-AUG-2000; 2000US-0225433P.

XX PA (RALP/) RALPH P.

XX PI Ralph P;

XX DR WPI; 2002-697507/75.

XX PT Determining an analyte in the presence of an interfering substance,  
XX PT comprises contacting a sample with an antibody recognizing analyte coated  
XX PT surface and interfering substance, and measuring unbound and bound  
XX PT analyte.

XX PS Example 1; Fig 1; 16pp; English.

XX CC The invention discloses a method for determining the amount of an analyte  
XX CC in a fluid sample in the presence of an interfering substance. The method  
XX CC comprises contacting a solid surface, dual-coated with an antibody  
XX CC recognising a free analyte and a second antibody recognising an  
XX CC interfering substance when bound to the analyte, with a fluid sample and  
XX CC then determining the total amount of free analyte and analyte bound to  
XX CC the interfering substance. The example in this specification discloses  
XX CC the quantitative determination of serum HERCEPTIN levels in patients  
XX CC undergoing HERCEPTIN therapy. The method accurately determines the amount  
XX CC of an anti-HER2 antibody in serum or plasma in the presence of an HER2  
XX CC (also referred to as ErbB2 and ErbB2 receptor, not defined) extracellular  
XX CC domain (ECD). The labelled secondary antibody recognises the anti-HER2  
XX CC antibody at an epitope different from that recognised by the first

CC antibody, as well as different from that recognised by the HER2 ECD. The  
CC method is useful for accurately determining the amount of an antibody,  
CC anti-HER2, that is circulating in a fluid sample (serum or plasma)  
CC derived from a cancer patient undergoing anti-HER2 therapy, in the  
CC presence of an interfering substance (the ECD) of the HER2 oncogene. The  
CC method prevents loss of detection in the qualitative assays and an  
CC underestimation in the quantitative assays. The sequence presented is the  
CC human HER2 receptor extracellular domain

XX SQ Sequence 645 AA;

Query Match 97.0%; Score 128; DB 5; Length 645;

Best Local Similarity 95.8%; Pred. No. 1.2e-10;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24

|||||

316 PLHNQEVTAEDGTQRAEKCKSPCA 339

Search completed: December 3, 2005, 07:43:42

Job time : 156.545 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:35:58 : Search time 35.4545 Seconds  
(without alignment)  
55.965 Million cell updates/sec

Title: US-09-632-036f-6

Perfect score: 132

Sequence: 1 PLHNQEVTAEDTQRAEKCCKPCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/H-COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/PCrUS-COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	419	2	US-09-630-155-2
2	128	97.0	624	2	US-08-422-108-1
3	128	97.0	624	2	US-08-422-734-1
4	128	97.0	645	2	US-09-602-812A-13
5	128	97.0	782	1	US-08-146-283-4
6	128	97.0	782	2	US-08-579-823A-4
7	128	97.0	782	2	US-09-344-195-4
8	128	97.0	1255	1	US-08-467-083-68
9	128	97.0	1255	1	US-08-414-417B-68
10	128	97.0	1255	1	US-08-484-438-8
11	128	97.0	1255	1	US-08-486-348A-68
12	128	97.0	1255	1	US-08-625-101-2
13	128	97.0	1255	1	US-08-625-101-2
14	128	97.0	1255	1	US-08-625-101-2
15	128	97.0	1255	2	US-08-356-786-2
16	128	97.0	1255	2	US-08-466-680B-68
17	128	97.0	1255	2	US-09-527-487-2
18	128	97.0	1255	2	US-09-811-115-3
19	128	97.0	1255	2	US-09-354-533-68
20	128	97.0	1255	2	US-09-441-411-6
21	71	53.8	15	1	US-09-167-516-2
22	71	53.8	15	1	US-08-457-083-56
23	71	53.8	15	1	US-08-414-417B-56
24	71	53.8	15	1	US-08-486-348A-56
25	71	53.8	15	1	US-08-468-545B-56
26	71	53.8	15	2	US-08-466-680B-56
27	63	47.7	18	1	US-09-354-533-56
					Sequence 60, Appl

28	63	47.7	18	1	US-08-414-417B-60	Sequence 60, Appl
29	63	47.7	18	1	US-08-486-348A-60	Sequence 60, Appl
30	63	47.7	18	1	US-08-488-545B-60	Sequence 60, Appl
31	63	47.7	18	2	US-08-466-680B-60	Sequence 60, Appl
32	63	47.7	18	2	US-09-354-533-60	Sequence 32, Appl
33	53	40.2	12	2	US-08-403-459-32	Sequence 3, Appl
34	52	39.4	273	1	US-09-055-095-3	Sequence 6, Appl
35	52	39.4	273	1	US-08-809-494A-6	Sequence 6, Appl
36	52	39.4	273	2	US-09-352-302-6	Sequence 6, Appl
37	52	39.4	746	2	US-09-949-016-11494	Sequence 11494, A
38	52	39.4	2324	2	US-09-902-540-9732	Sequence 9732, Ap
39	49	37.1	10	2	US-08-159-339A-285	Patent No. 5459061
40	49	37.1	76	6	5459061-1	Patent No. 5459061
41	49	37.1	76	6	5459061-10	Sequence 33572, A
42	49	37.1	151	2	US-09-270-767-32572	Sequence 47789, A
43	49	37.1	151	2	US-09-270-767-47789	Sequence 9, Appl
44	49	37.1	644	1	US-08-336-708A-9	Sequence 7, Appl
45	49	37.1	1210	1	US-08-484-438-7	Sequence 4, Appl
46	49	37.1	1210	1	US-08-475-035-4	Sequence 2, Appl
47	49	37.1	1210	2	US-09-715-249-2	Sequence 67, Appl
48	49	37.1	1210	2	US-09-723-307-67	Sequence 23999, A
49	48	36.4	260	2	US-09-252-991A-23999	Sequence 2, Appl
50	48	36.4	478	2	US-09-570-454-2	Sequence 2, Appl
51	48	36.4	478	2	US-09-867-521-2	Sequence 9675, Ap
52	47	35.6	61	2	US-09-489-039A-9675	Sequence 1, Appl
53	47	35.6	573	2	US-09-042-785A-2	Sequence 2, Appl
54	47	35.6	600	2	US-09-422-680A-10	Sequence 10, Appl
55	46	34.8	294	2	US-09-248-796A-25461	Sequence 25461, A
56	45.5	34.5	850	2	US-09-252-991A-20891	Sequence 20891, A
57	45	34.1	9	2	US-08-159-339A-282	Sequence 282, App
58	45	34.1	90	2	US-09-270-767-34477	Sequence 34477, A
59	45	34.1	90	2	US-09-270-767-49694	Sequence 49694, A
60	45	34.1	1106	2	US-09-949-016-9626	Sequence 9626, Ap
61	44.5	33.7	158	2	US-09-270-767-32962	Sequence 32962, A
62	44.5	33.7	158	2	US-09-270-767-48179	Sequence 48179, A
63	44.5	33.7	366	2	US-09-270-767-32253	Sequence 32253, A
64	44.5	33.7	366	2	US-09-270-767-47470	Sequence 47470, A
65	44.5	33.7	499	2	US-09-538-092-192	Sequence 192, App
66	44	33.3	76	2	US-09-631-816-29	Sequence 29, Appl
67	44	33.3	102	2	US-09-314-844F-6	Sequence 6, Appl
68	44	33.3	125	2	US-08-959-382-4	Sequence 4, Appl
69	44	33.3	125	2	US-09-314-844F-4	Sequence 4, Appl
70	44	33.3	147	2	US-09-527-236A-19	Sequence 19, Appl
71	44	33.3	147	2	US-09-756-854-19	Sequence 19, Appl
72	44	33.3	147	2	US-10-041-574-19	Sequence 19, Appl
73	44	33.3	147	2	US-09-095-094-19	Sequence 19, Appl
74	44	33.3	253	2	US-09-042-785A-4	Sequence 4, Appl
75	44	33.3	254	2	US-09-422-680A-6	Sequence 6, Appl
76	44	33.3	290	2	US-09-422-680A-2	Sequence 2, Appl
77	44	33.3	290	2	US-09-422-680A-8	Sequence 8, Appl
78	44	33.3	429	2	US-09-543-681A-4952	Sequence 4952, Ap
79	44	33.3	429	2	US-09-372-668-2	Sequence 2, Appl
80	44	33.3	429	2	US-09-697-377-2	Sequence 2, Appl
81	44	33.3	429	2	US-09-696-868-2	Sequence 2, Appl
82	44	33.3	429	2	US-09-697-341-2	Sequence 2, Appl
83	44	33.3	429	2	US-09-697-340-2	Sequence 2, Appl
84	44	33.3	429	2	US-10-115-195-2	Sequence 2, Appl
85	44	33.3	554	2	US-09-949-016-11730	Sequence 11730, A
86	44	33.3	605	2	US-09-042-785A-23	Sequence 23, Appl
87	44	33.3	609	2	US-09-396-149-9	Sequence 9, Appl
88	44	33.3	625	2	US-09-949-016-8500	Sequence 8500, Ap
89	44	33.3	655	2	US-08-959-382-2	Sequence 2, Appl
90	44	33.3	655	2	US-09-527-236A-2	Sequence 2, Appl
91	44	33.3	655	2	US-09-314-844F-2	Sequence 2, Appl
92	44	33.3	655	2	US-09-756-854-2	Sequence 2, Appl
93	44	33.3	655	2	US-09-999-833A-64	Sequence 64, Appl
94	44	33.3	655	2	US-10-041-574-2	Sequence 2, Appl
95	44	33.3	655	2	US-09-095-094-2	Sequence 2, Appl
96	44	33.3	655	2	US-10-020-445A-64	Sequence 64, Appl
97	44	33.3	691	2	US-09-422-680A-4	Sequence 4, Appl
98	43	32.6	119	2	US-10-101-464A-832	Sequence 832, App
99	43	32.6	120	2	US-09-228-986-128	Sequence 128, App
100	43	32.6	120	2	US-10-101-464A-128	Sequence 128, App

## ALIGNMENTS

## RESULT 1

US-09-630-155-2  
; Sequence 2, Application US/09630155  
; Patent No. 6414130  
; GENERAL INFORMATION:  
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVIS WRIGHT TREMAINE LLP  
; STREET: 1501 Fourth Avenue, 2600 Century Square  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: PC compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: Word

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/630,155  
FILING DATE: 16-Jan-2001  
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Davison, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206 628-7621  
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: polypeptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-630-155-2

Query Match 97.0%; Score 128; DB 2; Length 419;  
Best Local Similarity 95.8%; Pred. No. 2.2e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24

Db 316 PLHNQEVTAEDGTQRAEKCSPCA 339

## RESULT 2

US-08-422-108-1  
; Sequence 1, Application US/08422108  
; Patent No. 6015567  
; GENERAL INFORMATION:  
; APPLICANT: Hudziak, Robert M.  
; APPLICANT: Shepard, H. Michael  
; APPLICANT: Ullrich, Axel  
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-Apr-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319

FILING DATE: 19-MAY-1989

ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: 554C2D2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids

TYPE: Amino Acid  
TOPOLOGY: Linear

US-08-422-108-1

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24

Db 295 PLHNQEVTAEDGTQRAEKCSPCA 318

## RESULT 3

US-08-422-734-1  
; Sequence 1, Application US/08422734  
; Patent No. 6333169  
; GENERAL INFORMATION:  
; APPLICANT: Hudziak, Robert M.  
; APPLICANT: Shepard, H. Michael  
; APPLICANT: Ullrich, Axel  
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,734

FILING DATE: 435

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422108

FILING DATE: 14-Apr-1995

APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:







; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,417B  
; FILING DATE: 31-MAR-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-414-417B-68

Query Match 97.0%; Score 128; DB 1; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 7.6e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQVTAEDGTORAEKCKPC 24  
Db 316 PLHNQVTAEDGTORCEKCKPC 339

RESULT 10  
US-08-484-438-8  
; Sequence 8, Application US/08484438  
; Patent No. 5811098  
; Patent No. 5811098 5780031  
; GENERAL INFORMATION:  
; APPLICANT: Plowman, Gregory D.  
; APPLICANT: Culouscou, Jean-Michel  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Siegall, Clay B.  
; APPLICANT: Hellstr m, Ingegerd  
; APPLICANT: Hellstr m, Karl E.  
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,438  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/323,442  
; FILING DATE: 14-OCT-1994  
; APPLICATION NUMBER: US 08/150,704  
; FILING DATE: 10-NOV-1993

; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/981,165  
; FILING DATE: 24-NOV-1992  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-230  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-484-438-8

Query Match 97.0%; Score 128; DB 1; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 7.6e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQVTAEDGTORAEKCKPC 24  
Db 316 PLHNQVTAEDGTORCEKCKPC 339

RESULT 11  
US-08-486-348A-68  
; Sequence 68, Application US/08486348A  
; Patent No. 5846538  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,348A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-486-348A-68

Query Match 97.0%; Score 128; DB 1; Length 1255;

```
Best Local Similarity 95.8%; Pred. No. 7.6e-11; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 1;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24
    |||||
Db 316 PLHNQEVTAEDGTQCEKCSKPCA 339

RESULT 12
US-08-625-101-2
; Sequence 2, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-625-101-2

Query Match 97.0%; Score 128; DB 1; Length 1255;
Best Local Similarity 95.8%; Pred. No. 7.6e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24
    |||||
Db 316 PLHNQEVTAEDGTQCEKCSKPCA 339

RESULT 13
US-08-468-545B-68
; Sequence 68, Application US/08468545B
; Patent No. 5876712
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
```

```
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 97.0%; Score 128; DB 1; Length 1255;
Best Local Similarity 95.8%; Pred. No. 7.6e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24
    |||||
Db 316 PLHNQEVTAEDGTQCEKCSKPCA 339

RESULT 14
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 587305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
```



/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/354,533  
/ FILING DATE: 15-Jul-1999  
/ CLASSIFICATION: <Unknown>  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Sharkey, Richard G.  
/ REGISTRATION NUMBER: 32,629  
/ REFERENCE/DOCKET NUMBER: 920010.448C9  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (206) 622-4900  
/ TELEFAX: (206) 682-6031  
/ INFORMATION FOR SEQ ID NO: 68:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1255 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 68:  
US-09-354-533-68

Query Match 97.0%; Score 128; DB 2; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 7.6e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||  
DB 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

## RESULT 19

US-09-441-411-6  
/ Sequence 6, Application US/09441411  
/ Patent No. 6734172  
/ GENERAL INFORMATION:  
/ APPLICANT: Scholler, Nathalie B.  
/ APPLICANT: Disis, Mary L.  
/ APPLICANT: Hellstrom, Ingegerd  
/ APPLICANT: Hellstrom, Karl Erik  
/ TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
/ FILE REFERENCE: 730033.409  
/ CURRENT APPLICATION NUMBER: US/09/441,411  
/ CURRENT FILING DATE: 1999-11-16  
/ NUMBER OF SEQ ID NOS: 26  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 6  
/ LENGTH: 1255  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-09-441-411-6

Query Match 97.0%; Score 128; DB 2; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 7.6e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||  
DB 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

## RESULT 20

US-09-167-516-2  
/ Sequence 2, Application US/09167516  
/ Patent No. 6953573  
/ GENERAL INFORMATION:  
/ APPLICANT: Cheever, Martin A.  
/ APPLICANT: Disis, Mary L.  
/ TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
/ TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
/ TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
/ TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

/ NUMBER OF SEQUENCES: 4  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSES: SEED and BERRY LLP  
/ STREET: 6300 Columbia Center, 701 Fifth Avenue  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: USA  
/ ZIP: 98104-7092  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/167,516  
/ FILING DATE:

/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/625,101  
/ FILING DATE: 01-APR-1996  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Sharkey, Richard G.  
/ REGISTRATION NUMBER: 32,629  
/ REFERENCE/DOCKET NUMBER: 920010.448C7  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (206) 622-4900  
/ TELEFAX: (206) 682-6031  
/ INFORMATION FOR SEQ ID NO: 2:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1255 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-09-167-516-2

Query Match 97.0%; Score 128; DB 2; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 7.6e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||  
DB 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

## RESULT 21

US-08-467-083-56  
/ Sequence 56, Application US/08467083  
/ Patent No. 5726023  
/ GENERAL INFORMATION:  
/ APPLICANT: Cheever, Martin A.  
/ APPLICANT: Disis, Mary L.  
/ TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
/ TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
/ TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
/ NUMBER OF SEQUENCES: 68  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSES: Seed and Berry  
/ STREET: 6300 Columbia Center, 701 Fifth Avenue  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: US  
/ ZIP: 98104-7092  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/467,083  
/ FILING DATE: 06-JUN-1995  
/ CLASSIFICATION: 424  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/414,417

/ NUMBER OF SEQUENCES: 4  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSES: SEED and BERRY LLP  
/ STREET: 6300 Columbia Center, 701 Fifth Avenue  
/ CITY: Seattle  
/ STATE: Washington  
/ COUNTRY: USA  
/ ZIP: 98104-7092  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/467,083  
/ FILING DATE: 06-JUN-1995  
/ CLASSIFICATION: 424  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/414,417

; FILING DATE: 06-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; TELEX: 3723836 SEEDANBERRY  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-467-083-56

Query Match 53.8%; Score 71; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NQEVTAEDGTORAEK 18

Db 1 NQEVTAEDGTORCEK 15

## RESULT 22

US-08-414-417B-56  
; Sequence 56, Application US/08414417B  
; Patent No. 5801005

; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,417B  
; FILING DATE: 31-MAR-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear

US-08-414-417B-56

Query Match 53.8%; Score 71; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NQEVTAEDGTORAEK 18

Db 1 NQEVTAEDGTORCEK 15

## RESULT 23

US-08-486-348A-56  
; Sequence 56, Application US/08486348A  
; Patent No. 5846538

; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,348A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear

US-08-486-348A-56

Query Match 53.8%; Score 71; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NQEVTAEDGTORAEK 18

Db 1 NQEVTAEDGTORCEK 15

## RESULT 24

US-08-468-545B-56  
; Sequence 56, Application US/08468545B  
; Patent No. 5876712

; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092

Query Match 53.8%; Score 71; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NQEVTAEDGTORAEK 18

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,545B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C5  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear

US-08-468-545B-56

Query Match 53.8%; Score 71; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 NOEVTAEDGTQRAEK 18  
|||||  
DB 1 NOEVTAEDGTQRCCK 15  
|||||

RESULT 25  
US-08-466-680B-56  
; Sequence 56, Application US/08466680B  
; Patent No. 6075122  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/354,533  
; FILING DATE: 15-JUL-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: linear  
; TOPOLOGY: linear

US-09-354-533-56

Query Match 53.8%; Score 71; DB 2; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 NOEVTAEDGTQRAEK 18  
|||||  
DB 1 NOEVTAEDGTQRCCK 15  
|||||

RESULT 27  
US-08-467-083-60  
; Sequence 60, Application US/08467083  
; Patent No. 5726023  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; HER-2/neu ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,083  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/414,417  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEEDANBERRY  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-467-083-60

Query Match 47.7%; Score 63; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.0054;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 13 TORAEKCSKPCA 24  
Db 1 TORCEKCSKPCA 12

## RESULT 28

US-08-414-417B-60  
Sequence 60, Application US/08414417B  
Patent No. 5801005  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,417B  
FILING DATE: 31-MAR-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-414-417B-60

Query Match 47.7%; Score 63; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.0054;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 13 TORAEKCSKPCA 24  
Db 1 TORCEKCSKPCA 12

## RESULT 29

US-08-486-348A-60  
Sequence 60, Application US/08486348A  
Patent No. 5846538  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,348A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-486-348A-60

Query Match 47.7%; Score 63; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.0054;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 13 TORAEKCSKPCA 24  
Db 1 TORCEKCSKPCA 12

```
RESULT 30
US-08-468-545B-60
; Sequence 60, Application US/08468545B
; Patent No. 5876712
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C5
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-468-545B-60
Query Match 47.7%; Score 63; DB 1; Length 18;
Best Local Similarity 91.7%; Pred. No. 0.0054;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 TQRAEKCSKPCA 24
||| |||||
Db 1 TQCEKCSKPCA 12

RESULT 31
US-08-466-680B-60
; Sequence 60, Application US/0846680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-354-533-60
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-466-680B-60
Query Match 47.7%; Score 63; DB 2; Length 18;
Best Local Similarity 91.7%; Pred. No. 0.0054;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 TQRAEKCSKPCA 24
||| |||||
Db 1 TQCEKCSKPCA 12

RESULT 32
US-09-354-533-60
; Sequence 60, Application US/09354533
; Patent No. 6664370
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-354-533-60
```



Query Match 47.7%; Score 63; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 0.0054;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 13 TORAEKCKSPCA 24  
Db 1 TORCEKCKPCA 12

## RESULT 33

US-08-403-459-32  
; Sequence 32, Application US/08403459  
; Patent No. 6514942  
; GENERAL INFORMATION:  
; APPLICANT: Ioannides, Constantin G.  
; APPLICANT: Fisk, Bryan A.  
; APPLICANT: Ioannides, Maria G.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING  
; TITLE OF INVENTION: T-LYMPHOCYTES  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403.459  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitchell, Barbara S.  
; REGISTRATION NUMBER: 33,928  
; REFERENCE/DOCKET NUMBER: UTSC:390/KIT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-403-459-32

Query Match 40.2%; Score 53; DB 2; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NOEVTAEQTQR 15  
Db 1 NOEVTAWDTQR 12

## RESULT 34

US-09-055-095-3  
; Sequence 3, Application US/09055095  
; Patent No. 5945308  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Sather, Susan  
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/055,095  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0500 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 273 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1902984  
US-09-055-095-3

Query Match 39.4%; Score 52; DB 1; Length 273;  
Best Local Similarity 36.4%; Pred. No. 5.4;  
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 2 LHNQVTAEDGTORAEKCKSPC 23  
Db 123 LHHQNLNLQETLKRVCNCSAPC 144

## RESULT 35

US-08-809-494A-6  
; Sequence 6, Application US/0809494A  
; Patent No. 5962260  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya  
; APPLICANT: Masaki, Tomoo  
; TITLE OF INVENTION: Modified Low-Density Lipoprotein  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
; STREET: 261 Madison Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,494A  
; FILING DATE: 24-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: JP 6-321705
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-214206
; FILING DATE: 31-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldberg, Jules E
; REGISTRATION NUMBER: 24408
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 986-4090
; TELEFAX: 212 818-9479
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-494A-6

Query Match 39.4%; Score 52; DB 1; Length 273;
Best Local Similarity 36.4%; Pred. No. 5.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 LHNQEVTAEDGTQRAEKCCKPC 23
|||: : : : |||
Db 123 LHHQNLNLQETLKRVCNSAPC 144

RESULT 36
US-09-352-302-6
; Sequence 6, Application US/09352302
; Patent No. 6197937
; GENERAL INFORMATION:
; APPLICANT: Sawamura, Tatsuya
; APPLICANT: Masaki, Tomoo
; TITLE OF INVENTION: Modified Low-Density Lipoprotein
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel
; STREET: 261 Madison Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/352,302
; FILING DATE: 12-JUL-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-321705
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-214206
; FILING DATE: 31-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldberg, Jules E
; REGISTRATION NUMBER: 24408
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT/D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 986-4090
; TELEFAX: 212 818-9479
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; APPLICATION NUMBER: JP 6-321705
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-214206
; FILING DATE: 31-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldberg, Jules E
; REGISTRATION NUMBER: 24408
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 986-4090
; TELEFAX: 212 818-9479
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; US-09-352-302-6

Query Match 39.4%; Score 52; DB 2; Length 273;
Best Local Similarity 36.4%; Pred. No. 5.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 LHNQEVTAEDGTQRAEKCCKPC 23
|||: : : : |||
Db 123 LHHQNLNLQETLKRVCNSAPC 144

RESULT 37
US-09-949-016-11494
; Sequence 11494, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11494
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11494

Query Match 39.4%; Score 52; DB 2; Length 746;
Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCK 21
|||: : : : |||
Db 243 PLESQSAEGMSQRCCEGK 263

RESULT 38
US-09-902-540-9732
; Sequence 9732, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9732
; LENGTH: 2324
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-9732

Query Match 39.4%; Score 52; DB 2; Length 2324;
Best Local Similarity 45.5%; Pred. No. 60;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 HNQEVTAEDGTQRAEKCCKPCA 24
```

Db 494 HRRVVAADGAELAEERLKPAA 515

## RESULT 39

US-08-159-339A-285  
; Sequence 285, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Cells, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-0050300S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 285:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-285

Query Match 37.1%; Score 49; DB 2; Length 10;  
Best Local Similarity 90.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 GTQRAEKCSK 21  
Db 1 GTQCEKCSK 10

## RESULT 40

US-08-159-339A-285  
; Patent No. 5459061  
; APPLICANT: SATO, J DENRY;WU, DIANGING;WANG, LIHUA  
; TITLE OF INVENTION: HYBRIDOMAS PRODUCING MONOCLONAL ANTIBODIES  
; WHICH SPECIFICALLY BIND TO CONTINUOUS EPITOPE ON THE HUMAN EGF  
; RECEPTOR AND COMPETE WITH EGF FOR BINDING TO THE EGF RECEPTOR  
; NUMBER OF SEQUENCES: 10

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/133,274  
; FILING DATE: 07-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 470,642  
; FILING DATE: 26-JAN-1990  
; SEQ ID NO:1:  
; LENGTH: 76  
5459061-1

Query Match 37.1%; Score 49; DB 6; Length 76;  
Best Local Similarity 50.0%; Pred. No. 3.7;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 10 EDGTQRAEKCSKPC 23  
Db 2 EDGVRKCKCEGPC 15

Search completed: December 3, 2005, 07:49:48  
Job time : 37.4545 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:50:00 ; Search time 119.455 Seconds  
(without alignments)  
83.947 Million cell updates/sec

Title: US-09-632-036F-6  
Perfect score: 132  
Sequence: 1 PLHNOEVTARDGTQRAEKCKPCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
-Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications\_AA\_Main:  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	97.0	387	5	US-10-956-373-26
2	128	97.0	419	4	US-10-302-663-2
3	128	97.0	419	4	US-10-344-470-2
4	128	97.0	419	5	US-10-775-204-1811
5	128	97.0	419	5	US-10-846-113A-9
6	128	97.0	509	5	US-10-485-683-2
7	128	97.0	575	5	US-10-956-373-4
8	128	97.0	604	5	US-10-846-113A-7
9	128	97.0	607	6	US-11-021-951-183
10	128	97.0	614	5	US-10-956-373-16
11	128	97.0	633	5	US-10-846-113A-6
12	128	97.0	645	4	US-09-921-161-1
13	128	97.0	645	4	US-10-268-501-13
14	128	97.0	645	5	US-10-608-626-13
15	128	97.0	645	4	US-10-719-310-13
16	128	97.0	645	6	US-11-044-749-13
17	128	97.0	645	6	US-11-154-465-13
18	128	97.0	653	3	US-09-854-356-3
19	128	97.0	678	5	US-10-956-373-2
20	128	97.0	680	5	US-10-846-113A-5
21	128	97.0	685	4	US-10-412-804A-4
22	128	97.0	690	4	US-10-412-804A-11
23	128	97.0	712	3	US-09-854-356-7
24	128	97.0	715	4	US-10-412-804A-10
25	128	97.0	717	5	US-10-956-373-14
26	128	97.0	720	4	US-10-412-804A-6
27	128	97.0	919	3	US-09-854-356-6

28	128	97.0	1006	5	US-10-775-204-1643	Sequence 1643, Ap
29	128	97.0	1253	4	US-10-146-473-72	Sequence 72, Appli
30	128	97.0	1253	3	US-09-811-123-9	Sequence 9, Appli
31	128	97.0	1253	3	US-09-811-115-3	Sequence 3, Appli
32	128	97.0	1253	3	US-09-765-973-2	Sequence 2, Appli
33	128	97.0	1253	3	US-09-354-533-68	Sequence 68, Appli
34	128	97.0	1253	3	US-09-769-508-2	Sequence 2, Appli
35	128	97.0	1253	3	US-09-854-356-1	Sequence 1, Appli
36	128	97.0	1253	3	US-09-930-125-2	Sequence 2, Appli
37	128	97.0	1253	3	US-09-441-411-6	Sequence 6, Appli
38	128	97.0	1253	3	US-09-984-092-4	Sequence 4, Appli
39	128	97.0	1253	4	US-10-207-655-45	Sequence 45, Appli
40	128	97.0	1253	4	US-10-177-293-126	Sequence 126, App
41	128	97.0	1253	4	US-10-207-498-6	Sequence 6, Appli
42	128	97.0	1253	4	US-10-338-730-2	Sequence 2, Appli
43	128	97.0	1253	4	US-10-313-644-2	Sequence 2, Appli
44	128	97.0	1253	4	US-10-322-892-4	Sequence 4, Appli
45	128	97.0	1253	4	US-10-272-437A-28	Sequence 28, Appli
46	128	97.0	1253	4	US-10-117-937-594	Sequence 594, App
47	128	97.0	1253	4	US-10-418-027-3	Sequence 3, Appli
48	128	97.0	1253	4	US-10-394-322A-17	Sequence 17, Appli
49	128	97.0	1253	4	US-10-345-871-553	Sequence 553, App
50	128	97.0	1253	4	US-10-435-696-36	Sequence 36, Appli
51	128	97.0	1253	4	US-10-149-138-4641	Sequence 4641, Ap
52	128	97.0	1253	4	US-10-469-162-3	Sequence 3, Appli
53	128	97.0	1253	4	US-10-253-286-553	Sequence 553, App
54	128	97.0	1253	4	US-10-647-005-68	Sequence 68, Appli
55	128	97.0	1253	4	US-10-149-138-4641	Sequence 4641, Ap
56	128	97.0	1253	4	US-10-441-779C-4	Sequence 4, Appli
57	128	97.0	1253	4	US-10-734-564-126	Sequence 126, App
58	128	97.0	1253	4	US-10-657-022-90	Sequence 90, Appli
59	128	97.0	1253	5	US-10-762-128-6	Sequence 6, Appli
60	128	97.0	1253	5	US-10-484-067-1	Sequence 1, Appli
61	128	97.0	1253	5	US-10-723-860-9	Sequence 9, Appli
62	128	97.0	1253	5	US-10-615-343-17	Sequence 17, Appli
63	128	97.0	1253	5	US-10-794-514A-1	Sequence 1, Appli
64	128	97.0	1253	5	US-10-871-708-9	Sequence 9, Appli
65	128	97.0	1253	5	US-10-983-340-17	Sequence 17, Appli
66	128	97.0	1253	5	US-10-846-113A-27	Sequence 27, Appli
67	128	97.0	1253	6	US-11-067-064-594	Sequence 594, App
68	128	97.0	1253	6	US-11-121-347-68	Sequence 68, Appli
69	128	97.0	1253	6	US-11-037-713-12	Sequence 12, Appli
70	128	97.0	1253	6	US-11-067-159-594	Sequence 594, App
71	128	97.0	1294	5	US-10-956-373-10	Sequence 10, Appli
72	124	93.9	23	6	US-11-009-533-50	Sequence 50, Appli
73	114	86.4	139	5	US-10-949-667-4	Sequence 4, Appli
74	114	86.4	654	3	US-09-854-356-8	Sequence 8, Appli
75	114	86.4	1256	3	US-09-854-356-2	Sequence 2, Appli
76	114	86.4	1256	3	US-09-854-356-14	Sequence 14, Appli
77	114	86.4	1257	5	US-10-484-067-2	Sequence 2, Appli
78	114	86.4	1260	3	US-09-870-759-118	Sequence 118, App
79	114	86.4	1260	3	US-09-751-708A-118	Sequence 118, App
80	114	86.4	1260	4	US-10-428-817A-114	Sequence 114, App
81	114	86.4	1260	5	US-10-937-758A-95	Sequence 95, Appli
82	114	86.4	1260	5	US-10-949-667-8	Sequence 8, Appli
83	71	53.8	15	3	US-09-354-533-56	Sequence 56, Appli
84	71	53.8	15	4	US-10-149-138-3929	Sequence 3929, Ap
85	71	53.8	15	4	US-10-149-138-4517	Sequence 4517, Ap
86	71	53.8	15	4	US-10-847-005-56	Sequence 56, Appli
87	71	53.8	15	4	US-10-149-138-3929	Sequence 3929, Ap
88	71	53.8	15	4	US-10-149-138-4517	Sequence 4517, Ap
89	71	53.8	15	6	US-11-121-347-56	Sequence 56, Appli
90	70	53.0	15	4	US-10-149-138-3891	Sequence 3891, Ap
91	70	53.0	15	4	US-10-149-138-4457	Sequence 4457, Ap
92	70	53.0	15	4	US-10-149-138-3891	Sequence 3891, Ap
93	70	53.0	15	4	US-10-149-138-4457	Sequence 4457, Ap
94	63	47.7	18	3	US-09-354-533-60	Sequence 60, Appli
95	63	47.7	18	4	US-10-245-871-578	Sequence 578, App
96	63	47.7	18	4	US-10-253-286-578	Sequence 578, App
97	63	47.7	18	4	US-10-647-005-60	Sequence 60, Appli
98	63	47.7	18	6	US-11-121-347-60	Sequence 60, Appli
99	55	41.7	11	4	US-10-149-138-2532	Sequence 2532, Ap
100	55	41.7	11	4	US-10-149-138-3237	Sequence 3237, Ap

ALIGNMENTS

```
RESULT 1
US-10-956-373-26
; Sequence 26, Application US/10956373
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
; APPLICANT: Oren, Anat
; APPLICANT: Rotman, Galit
; APPLICANT: Sela-Tavor, Osnat
; APPLICANT: Walach, Shira
; APPLICANT: Sameah-Greenwald, Shirley
; APPLICANT: Beiman, Merav
; APPLICANT: Eshel, Dani
; APPLICANT: Savitsky, Kinneret
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND
; FILE REFERENCE: 28399
; CURRENT APPLICATION NUMBER: US/10/956,373
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-956-373-26

Query Match          97.0%; Score 128; DB 5; Length 387;
Best Local Similarity 95.8%; Pred. No. 4.9e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
| | | | | | | | | | | | | | | | | | | | | |
DB 316 PLHNQEVTAEDGTQRCCKSKPKCA 339

RESULT 2
US-10-302-663-2
; Sequence 2, Application US/10302663
; Publication No. US20040022785A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE HER-2/NEU PRODUCT, IN CEL
; TITLE OF INVENTION: EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECEPTOR A
; FILE REFERENCE: 49321-73
; CURRENT APPLICATION NUMBER: US/10/302,663
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 342
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 345
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 346
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; FEATURE:
```

```
; NAME/KEY: VARIANT
; LOCATION: 356
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 358
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 361
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variant
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 376
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 394
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 404
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 413
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
US-10-302-663-2

Query Match          97.0%; Score 128; DB 4; Length 419;
Best Local Similarity 95.8%; Pred. No. 5.3e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
| | | | | | | | | | | | | | | | | | | | | |
DB 316 PLHNQEVTAEDGTQRCCKSKPKCA 339

RESULT 3
US-10-344-470-2
; Sequence 2, Application US/10344470
; Publication No. US20040052796A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN
; TITLE OF INVENTION: THAT EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECE
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: 49321-81
; CURRENT APPLICATION NUMBER: US/10/344,470
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT / US01/25502
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (342)..(342)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

LOCATION: (346)..(346)  
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (356)..(356)  
OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (358)..(358)  
OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (361)..(361)  
OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variants  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (376)..(376)  
OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (394)..(394)  
OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (404)..(404)  
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (413)..(413)  
OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this  
OTHER INFORMATION: position  
US-10-344-470-2

Query Match 97.0%; Score 128; DB 4; Length 419;

Best Local Similarity 95.8%; Pred. No. 5.3e-10;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||

Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

#### RESULT 4

US-10-775-204-1811

Sequence 1811, Application US/10775204

Publication No. US20050186664A1

GENERAL INFORMATION:

APPLICANT: Rosen, Craig A.

APPLICANT: Haseltine, William A.

APPLICANT: Balance, David J.

APPLICANT: Turner, Andrew J.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF564

CURRENT APPLICATION NUMBER: US/10775,204

CURRENT FILING DATE: 2004-02-11

PRIOR APPLICATION NUMBER: 60/341,811

PRIOR FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 60/360,000

PRIOR FILING DATE: 2002-02-28

PRIOR APPLICATION NUMBER: 60/378,950

PRIOR FILING DATE: 2002-05-10

PRIOR APPLICATION NUMBER: 60/398,008

PRIOR FILING DATE: 2002-07-24

PRIOR APPLICATION NUMBER: 60/411,355

PRIOR FILING DATE: 2002-09-18

PRIOR APPLICATION NUMBER: 60/414,984

PRIOR FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: 60/417,611  
PRIOR FILING DATE: 2002-10-11  
PRIOR APPLICATION NUMBER: 60/420,246  
PRIOR FILING DATE: 2002-10-23  
PRIOR APPLICATION NUMBER: 60/423,623  
PRIOR FILING DATE: 2002-11-05  
PRIOR APPLICATION NUMBER: 60/351,360  
PRIOR FILING DATE: 2002-01-28  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 2222  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1811  
LENGTH: 419  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-775-204-1811

Query Match 97.0%; Score 128; DB 5; Length 419;

Best Local Similarity 95.8%; Pred. No. 5.3e-10;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||

Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

#### RESULT 5

US-10-846-113A-9

Sequence 9, Application US/10846113A

Publication No. US2005023908A1

GENERAL INFORMATION:

APPLICANT: Shepard, H. Michael

APPLICANT: Clinton, Gail M.

APPLICANT: Lackey, David B.

APPLICANT: Jin, Pei

TITLE OF INVENTION: INTRON FUSION PROTEINS, AND

TITLE OF INVENTION: METHODS OF IDENTIFYING AND USING SAME

FILE REFERENCE: 17118-006001

CURRENT APPLICATION NUMBER: US/10/846,113A

CURRENT FILING DATE: 2004-05-14

PRIOR APPLICATION NUMBER: US 60/471,141

PRIOR FILING DATE: 2003-05-16

NUMBER OF SEQ ID NOS: 53

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 419

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Human ERBB2.c BUILD 31

US-10-846-113A-9

Query Match

Best Local Similarity 97.0%; Score 128; DB 5; Length 419;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||

Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

#### RESULT 6

US-10-485-683-2

Sequence 2, Application US/10485683

Publication No. US20040248196A1

GENERAL INFORMATION:

APPLICANT: ADAMS, TIMOTHY EDWARD

APPLICANT: BURGESS, ANTHONY WILKS

APPLICANT: ELLEMAN, THOMAS CHARLES

APPLICANT: GARRETT, THOMAS PETER JOHN

APPLICANT: JORISSEN, ROBERT NICHOLAS

APPLICANT: LOU, MEIZHEN

; APPLICANT: LOVREZCZ, GEORGE OSCAR  
; APPLICANT: MCKERN, NEIL MORETON  
; APPLICANT: NICE, EDOUARD COLLINS  
; APPLICANT: WARD, COLIN WESLEY  
; TITLE OF INVENTION: METHODS OF SCREENING BASED ON THE EGF RECEPTOR CRYSTAL  
; TITLE OF INVENTION: STRUCTURE  
; FILE REFERENCE: 051654/0104  
; CURRENT APPLICATION NUMBER: US/10/485,683  
; CURRENT FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: PCT/AU02/01042  
; PRIOR FILING DATE: 2002-08-05  
; PRIOR APPLICATION NUMBER: AU PR 6827  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: AU PR 6828  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: 60/336,560  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/335,393  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: AU PS 2731  
; PRIOR FILING DATE: 2002-05-31  
; PRIOR APPLICATION NUMBER: 60/388,171  
; PRIOR FILING DATE: 2002-06-11  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 2  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-485-683-2

Query Match 97.0%; Score 128; DB 5; Length 509;  
Best Local Similarity 95.8%; Pred. No. 6.5e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRCCKSKPCA 24  
|||||

Db 295 PLHNQEVTAEDGTQRCCKSKPCA 318  
|||||

RESULT 7  
US-10-956-373-4  
; Sequence 4, Application US/10956373  
; Publication No. US20050123538A1  
; GENERAL INFORMATION:  
; APPLICANT: Shemesh, Ronen  
; APPLICANT: Oren, Anat  
; APPLICANT: Rotman, Galit  
; APPLICANT: Sela-Tavor, Osnat  
; APPLICANT: Walach, Shira  
; APPLICANT: Sameah-Greenwald, Shirley  
; APPLICANT: Beiman, Merav  
; APPLICANT: Eshel, Dani  
; APPLICANT: Savitsky, Kinneret  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND  
; TITLE OF INVENTION: METHODS USING SAME  
; FILE REFERENCE: 28399  
; CURRENT APPLICATION NUMBER: US/10/956,373  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-956-373-4

Query Match 97.0%; Score 128; DB 5; Length 575;  
Best Local Similarity 95.8%; Pred. No. 7.5e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRCCKSKPCA 24  
|||||

Db 316 PLHNQEVTAEDGTQRCCKSKPCA 339

## RESULT 8

US-10-846-113A-7  
; Sequence 7, Application US/10846113A  
; Publication No. US20050239088A1  
; GENERAL INFORMATION:  
; APPLICANT: Shepard, H. Michael  
; APPLICANT: Clinton, Gail M.  
; APPLICANT: Lackey, David B.  
; APPLICANT: Jin, Pei  
; TITLE OF INVENTION: INTRON FUSION PROTEINS, AND  
; TITLE OF INVENTION: METHODS OF IDENTIFYING AND USING SAME  
; FILE REFERENCE: 17118-006001  
; CURRENT APPLICATION NUMBER: US/10/846,113A  
; CURRENT FILING DATE: 2004-05-14  
; PRIOR APPLICATION NUMBER: US 60/471,141  
; PRIOR FILING DATE: 2003-05-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human ERBB2.1.f  
; US-10-846-113A-7

Query Match 97.0%; Score 128; DB 5; Length 575;  
Best Local Similarity 95.8%; Pred. No. 7.5e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRCCKSKPCA 24  
|||||

Db 316 PLHNQEVTAEDGTQRCCKSKPCA 339

## RESULT 9

US-11-021-951-183  
; Sequence 183, Application US/11021951  
; Publication No. US20050175581A1  
; GENERAL INFORMATION:  
; APPLICANT: HAUPTS, Ulrich  
; APPLICANT: KOLTERMANN, Andre  
; APPLICANT: SCHEIDIG, Andreas  
; APPLICANT: VOTSMEIER, Christian  
; APPLICANT: Kettling, Ulrich  
; APPLICANT: COCO, Wayne Michael  
; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical  
; TITLE OF INVENTION: And Diagnostic Use Thereof  
; FILE REFERENCE: 04156.000205  
; CURRENT APPLICATION NUMBER: US/11/021,951  
; CURRENT FILING DATE: 2004-12-22  
; PRIOR APPLICATION NUMBER: 10/872,198  
; PRIOR FILING DATE: 2004-06-18  
; PRIOR APPLICATION NUMBER: 60/543,518  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/524,960  
; PRIOR FILING DATE: 2003-11-25  
; PRIOR APPLICATION NUMBER: EP 04003058  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: EP 03025871  
; PRIOR FILING DATE: 2003-11-11  
; PRIOR APPLICATION NUMBER: EP 03025851  
; PRIOR FILING DATE: 2003-11-10  
; PRIOR APPLICATION NUMBER: EP 03013819  
; PRIOR FILING DATE: 2003-06-18  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 183  
; LENGTH: 607  
; TYPE: PRT



```
; ORGANISM: Homo sapiens
US-11-021-951-183

Query Match          97.0%; Score 128; DB 6; Length 607;
Best Local Similarity 95.8%; Pred. No. 7.9e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
   |||||
Db 294 PLHNQEVTAEDGTQRAEKCKSPCA 317

RESULT 10
US-10-956-373-16
; Sequence 16, Application US/10956373
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
; APPLICANT: Oren, Anat
; APPLICANT: Rotman, Galit
; APPLICANT: Sela-Tavor, Osnat
; APPLICANT: Walach, Shira
; APPLICANT: Sameah-Greenwald, Shirley
; APPLICANT: Beiman, Merav
; APPLICANT: Eshel, Dani
; APPLICANT: Savitsky, Kinneret
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND
; FILE REFERENCE: 28399
; CURRENT APPLICATION NUMBER: US/10/956,373
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-956-373-16

Query Match          97.0%; Score 128; DB 5; Length 614;
Best Local Similarity 95.8%; Pred. No. 8e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

RESULT 11
US-10-846-113A-6
; Sequence 6, Application US/10846113A
; Publication No. US2005023908A1
; GENERAL INFORMATION:
; APPLICANT: Shepard, H. Michael
; APPLICANT: Clinton, Gail M.
; APPLICANT: Lackey, David B.
; APPLICANT: Jin, Pei
; TITLE OF INVENTION: INTRON FUSION PROTEINS, AND
; FILE REFERENCE: 17118-006001
; CURRENT APPLICATION NUMBER: US/10/846,113A
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US 60/471,141
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human ERBB2.1.e
US-10-846-113A-6
```

```
Query Match          97.0%; Score 128; DB 5; Length 633;
Best Local Similarity 95.8%; Pred. No. 8.3e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

RESULT 12
US-09-921-161-1
; Sequence 1, Application US/09921161
; Patent No. US20020090662A1
; GENERAL INFORMATION:
; APPLICANT: Ralph, Peter
; TITLE OF INVENTION: ANALYTICAL METHOD
; FILE REFERENCE: GENENT.066A
; CURRENT APPLICATION NUMBER: US/09/921,161
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/225,433
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-161-1

Query Match          97.0%; Score 128; DB 3; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

RESULT 13
US-10-268-501-13
; Sequence 13, Application US/10268501
; Publication No. US20030086924A1
; GENERAL INFORMATION:
; APPLICANT: Slikowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: PI467R2P1
; CURRENT APPLICATION NUMBER: US/10/268,501
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 05/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-268-501-13

Query Match          97.0%; Score 128; DB 4; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
   |||||
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

RESULT 14
US-10-608-626-13
; Sequence 13, Application US/10608626
; Publication No. US20040013667A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Kelsey, Stephen M.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2P2
; CURRENT APPLICATION NUMBER: US/10/608,626
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-608-626-13

Query Match          97.0%; Score 128; DB 4; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339

RESULT 15
US-10-719-310-13
; Sequence 13, Application US/10719310
; Publication No. US20040258685A1
; GENERAL INFORMATION:
; APPLICANT: Brunetta, Paul G.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: THERAPY OF NON-MALIGNANT DISEASES OR DISORDERS WITH
; TITLE OF INVENTION: ANTI-ERBB2 ANTIBODIES
; FILE REFERENCE: P1979R1
; CURRENT APPLICATION NUMBER: US/10/719,310
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/428,027
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: human
US-10-719-310-13

Query Match          97.0%; Score 128; DB 5; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339

RESULT 16
US-11-044-749-13
; Sequence 13, Application US/11044749
; Publication No. US20050208043A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Camellia W.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Humanized Anti-ErbB2 Antibodies and Treatment with
; TITLE OF INVENTION: Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2
; CURRENT APPLICATION NUMBER: US/11/044,749
; CURRENT FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US/09/602,812

; GENERAL INFORMATION:
; APPLICANT: Kelsey, Stephen M.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2P2
; CURRENT APPLICATION NUMBER: US/10/608,626
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-044-749-13

Query Match          97.0%; Score 128; DB 6; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339

RESULT 17
US-11-154-465-13
; Sequence 13, Application US/11154465
; Publication No. US20050238640A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Camellia W.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Humanized Anti-ErbB2 Antibodies and Treatment with
; TITLE OF INVENTION: Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2
; CURRENT APPLICATION NUMBER: US/11/154,465
; CURRENT FILING DATE: 2005-06-16
; PRIOR APPLICATION NUMBER: US/09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-154-465-13

Query Match          97.0%; Score 128; DB 6; Length 645;
Best Local Similarity 95.8%; Pred. No. 8.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24
Db 316 PLHNQEVTAEDGTQRAEKCKPCA 339

RESULT 18
US-09-854-356-3
; Sequence 3, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheever, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 653
```

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu  
US-09-854-356-3

Query Match 97.0%; Score 128; DB 3; Length 653;  
Best Local Similarity 95.8%; Pred. No. 8.6e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

## RESULT 19

US-10-956-373-2  
; Sequence 2, Application US/10956373  
; Publication No. US20050123538A1  
; GENERAL INFORMATION:  
; APPLICANT: Shemesh, Ronen  
; APPLICANT: Oren, Anat  
; APPLICANT: Rotman, Galit  
; APPLICANT: Sela-Tavor, Osnat  
; APPLICANT: Walach, Shira  
; APPLICANT: Sameah-Greenwald, Shirley  
; APPLICANT: Beiman, Merav  
; APPLICANT: Eshel, Dani  
; APPLICANT: Savitsky, Kinneret  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ErbB-2 POLYPEPTIDES AND KITS AND  
; TITLE OF INVENTION: METHODS USING SAME  
; FILE REFERENCE: 28399  
; CURRENT APPLICATION NUMBER: US/10/956,373  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-956-373-2

Query Match 97.0%; Score 128; DB 5; Length 678;  
Best Local Similarity 95.8%; Pred. No. 8.9e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

## RESULT 20

US-10-846-113A-5  
; Sequence 5, Application US/10846113A  
; Publication No. US20050239088A1  
; GENERAL INFORMATION:  
; APPLICANT: Shepard, H. Michael  
; APPLICANT: Clinton, Gail M.  
; APPLICANT: Lackey, David B.  
; APPLICANT: Jin, Pei  
; TITLE OF INVENTION: INTRON FUSION PROTEINS, AND  
; TITLE OF INVENTION: METHODS OF IDENTIFYING AND USING SAME  
; FILE REFERENCE: 17118-006001  
; CURRENT APPLICATION NUMBER: US/10/846,113A  
; CURRENT FILING DATE: 2004-05-14  
; PRIOR FILING DATE: 2004-05-14  
; PRIOR FILING DATE: 2003-05-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 680  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Human ERBB2.1.d  
US-10-846-113A-5

Query Match 97.0%; Score 128; DB 5; Length 680;  
Best Local Similarity 95.8%; Pred. No. 8.9e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
|||||  
Db 318 PLHNQEVTAEDGTQRAEKCKSPCA 341

## RESULT 21

US-10-412-804A-4  
; Sequence 4, Application US/10412804A  
; Publication No. US20030228606A1  
; GENERAL INFORMATION:  
; APPLICANT: Jing, Shuguan  
; APPLICANT: Tatarewicz, Suzanna  
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 01-1624-A  
; CURRENT APPLICATION NUMBER: US/10/412,804A  
; CURRENT FILING DATE: 2003-04-11  
; PRIOR APPLICATION NUMBER: 60/371,912  
; PRIOR FILING DATE: 2002-04-11  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-412-804A-4

Query Match 97.0%; Score 128; DB 4; Length 685;  
Best Local Similarity 95.8%; Pred. No. 9e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

## RESULT 22

US-10-412-804A-11  
; Sequence 11, Application US/10412804A  
; Publication No. US20030228606A1  
; GENERAL INFORMATION:  
; APPLICANT: Jing, Shuguan  
; APPLICANT: Tatarewicz, Suzanna  
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 01-1624-A  
; CURRENT APPLICATION NUMBER: US/10/412,804A  
; CURRENT FILING DATE: 2003-04-11  
; PRIOR APPLICATION NUMBER: 60/371,912  
; PRIOR FILING DATE: 2002-04-11  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-412-804A-11

Query Match 97.0%; Score 128; DB 4; Length 690;  
Best Local Similarity 95.8%; Pred. No. 9.1e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCKSPCA 339

```
RESULT 23
US-09-854-356-7
; Sequence 7, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheever, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US 09 854 356 7

Query Match          97.0%; Score 128; DB 3; Length 712;
Best Local Similarity 95.8%; Pred. No. 9.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTTORAEKCKSPCA 24
DB 316 PLHNOEVTADGTTORAEKCKSPCA 339

* RESULT 24
US-10-412-804A-10
; Sequence 10, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-10

Query Match          97.0%; Score 128; DB 4; Length 715;
Best Local Similarity 95.8%; Pred. No. 9.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTTORAEKCKSPCA 24
DB 316 PLHNOEVTADGTTORAEKCKSPCA 339

RESULT 25
US-10-956-373-14
; Sequence 14, Application US/10956373
```

```
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
; APPLICANT: Oren, Anat
; APPLICANT: Rotman, Galit
; APPLICANT: Seia-Tavor, Osnat
; APPLICANT: Walach, Shira
; APPLICANT: Sameah-Greenwald, Shirley
; APPLICANT: Beiman, Merav
; APPLICANT: Eshel, Dani
; APPLICANT: Savitsky, Kinneret
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ErbB-2 POLYPEPTIDES AND KITS AND
; TITLE OF INVENTION: METHODS USING SAME
; FILE REFERENCE: 28399
; CURRENT APPLICATION NUMBER: US/10/956,373
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-956-373-14

Query Match          97.0%; Score 128; DB 5; Length 717;
Best Local Similarity 95.8%; Pred. No. 9.5e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTTORAEKCKSPCA 24
DB 316 PLHNOEVTADGTTORAEKCKSPCA 339

RESULT 26
US-10-412-804A-6
; Sequence 6, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-6

Query Match          97.0%; Score 128; DB 4; Length 720;
Best Local Similarity 95.8%; Pred. No. 9.5e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTTORAEKCKSPCA 24
DB 316 PLHNOEVTADGTTORAEKCKSPCA 339

RESULT 27
US-09-854-356-6
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheever, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
```

```
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match          97.0%; Score 128; DB 3; Length 919;
Best Local Similarity 95.8%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

-QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
    |||||
Db 316 PLHNQEVTAEDGTQRCCKCKPKCA 339

* RESULT 28
US-10-775-204-1643
; Sequence 1643, Application US/10775204
; Publication No. US20050186664A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; APPLICANT: Balance, David J.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF564
; CURRENT APPLICATION NUMBER: US/10/775,204
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/341,811
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/360,000
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/378,950
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/398,008
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/411,355
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/414,984
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/417,611
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/420,246
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 60/423,623
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 60/351,360
; PRIOR FILING DATE: 2002-01-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2222
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1643
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-775-204-1643

Query Match          97.0%; Score 128; DB 5; Length 1006;
Best Local Similarity 95.8%; Pred. No. 1.4e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match          97.0%; Score 128; DB 3; Length 919;
Best Local Similarity 95.8%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

-QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
    |||||
Db 316 PLHNQEVTAEDGTQRCCKCKPKCA 339

* RESULT 29
US-10-146-473-72
; Sequence 72, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-72

Query Match          97.0%; Score 128; DB 4; Length 1253;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

-QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
    |||||
Db 316 PLHNQEVTAEDGTQRCCKCKPKCA 339

* RESULT 30
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB
; FILE REFERENCE: ANTIBODY-MAYTANSINOID CONJUGATES
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9

Query Match          97.0%; Score 128; DB 3; Length 1255;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

-QY 1 PLHNQEVTAEDGTQRAEKCKPKCA 24
    |||||
Db 316 PLHNQEVTAEDGTQRCCKCKPKCA 339
```

```
RESULT 31
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: King, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match          97.0%; Score 128; DB 3; Length 1255;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTQRAEKCSPCA 24
DB 316 PLHNOEVTADGTQRCCKSPCA 339

RESULT 32
US-09-765-973-2
; Sequence 2, Application US/09765973
; Publication No. US20020039573A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmermann, Susan
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND
; TREATMENT OF HER-2/neu ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.496
; CURRENT APPLICATION NUMBER: US/09/765,973
; CURRENT FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-765-973-2

Query Match          97.0%; Score 128; DB 3; Length 1255;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTQRAEKCSPCA 24
DB 316 PLHNOEVTADGTQRCCKSPCA 339

RESULT 33
US-09-354-533-68
; Sequence 68, Application US/09354533
; Publication No. US20020055614A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-354-533-68

Query Match          97.0%; Score 128; DB 3; Length 1255;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTQRAEKCSPCA 24
DB 316 PLHNOEVTADGTQRCCKSPCA 339

RESULT 34
US-09-769-508-2
; Sequence 2, Application US/09769508
; Patent No. US2002015527A1
; GENERAL INFORMATION:
; APPLICANT: STUART, SUSAN G.
; APPLICANT: MONAHAN, JOHN J.
; APPLICANT: LANGTON, BEATRICE CLAUDIA
; APPLICANT: HANCOCK, MIRIAM E.C.
; APPLICANT: CHAO, LORRINE A.
; APPLICANT: BLUFORD, PETER
; TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
; FILE REFERENCE: BBIO-111-C1
; CURRENT APPLICATION NUMBER: US/09/769,508
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

Query Match          97.0%; Score 128; DB 3; Length 1255;
Best Local Similarity 95.8%; Pred. No. 1.7e-09;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNOEVTADGTQRAEKCSPCA 24
DB 316 PLHNOEVTADGTQRCCKSPCA 339

RESULT 35
US-09-854-356-1
```

; Sequence 1, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Cheyssen, Dirk  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: SmithKline Beecham Biologicals S. A.  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human HER-2/neu protein  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(653)  
; OTHER INFORMATION: extracellular domain (ECD)  
; NAME/KEY: DOMAIN  
; LOCATION: (676)..(1255)  
; OTHER INFORMATION: intracellular domain (ICD)  
; NAME/KEY: DOMAIN  
; LOCATION: (990)..(1255)  
; OTHER INFORMATION: phosphorylation domain (PD)  
; NAME/KEY: DOMAIN  
; LOCATION: (990)..(1048)  
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred  
; OTHER INFORMATION: portion (delta PD)  
US-09-854-356-1

Query Match 97.0%; Score 128; DB 3; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 36  
US-09-930-125-2  
; Sequence 2, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
; FILE REFERENCE: 210121-544  
; CURRENT APPLICATION NUMBER: US/09/930,125  
; CURRENT FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-930-125-2

Query Match 97.0%; Score 128; DB 3; Length 1255;

Best Local Similarity 95.8%; Pred. No. 1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 37  
US-09-441-411-6  
; Sequence 6, Application US/09441411  
; Publication No. US20030008342A1  
; GENERAL INFORMATION:  
; APPLICANT: Scholler, Nathalie B.  
; APPLICANT: Disis, Mary L.  
; APPLICANT: Hellstrom, Ingegerd  
; APPLICANT: Hellstrom, Karl Erik  
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
; FILE REFERENCE: 730033.409  
; CURRENT APPLICATION NUMBER: US/09/441,411  
; CURRENT FILING DATE: 1999-11-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-441-411-6

Query Match 97.0%; Score 128; DB 3; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 38  
US-09-984-092-4  
; Sequence 4, Application US/09984092  
; Publication No. US20040037840A1  
; GENERAL INFORMATION:  
; APPLICANT: Pharmexa A/S  
; TITLE OF INVENTION: NOVEL THERAPEUTIC VACCINE FORMULATIONS  
; FILE REFERENCE: FI011PC00  
; CURRENT APPLICATION NUMBER: US/09/984,092  
; CURRENT FILING DATE: 2001-10-26  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-092-4

Query Match 97.0%; Score 128; DB 3; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCCKPCA 24  
|||||  
Db 316 PLHNQEVTAEDGTQRAEKCCKPCA 339

RESULT 39  
US-10-207-655-45  
; Sequence 45, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS

; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 45  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-207-655-45

Query Match 97.0%; Score 128; DB 4; Length 1255;  
Best Local Similarity 95.8%; Pred.No.1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24  
DB 316 PLHNQEVTAEDGTQRAEKCSPCA 339

RESULT 40  
US-10-177-293-126  
; Sequence 126, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Gannavarpu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Pusztai, Lajos  
; APPLICANT: Meric, Funda  
; APPLICANT: Sahin, Aysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-038  
; CURRENT APPLICATION NUMBER: US/10/177,293  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299,887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301,572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306,501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325,002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362,585  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-177-293-126

Query Match 97.0%; Score 128; DB 4; Length 1255;  
Best Local Similarity 95.8%; Pred.No.1.7e-09;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCSPCA 24

DB 316 PLHNQEVTAEDGTQRAEKCSPCA 339  
Search completed: December 3, 2005, 08:08:27  
Job time : 122.455 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 3, 2005, 07:48:55 ; Search time 7.09091 Seconds  
(without alignments)  
16.207 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLHNQEVTAEDGTQRAEKSKPCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

## Database :

Published Applications\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	128	97.0	1255	7	US-11-022-562-213
2	80	60.6	17	7	US-11-060-646-8
3	39	29.5	52	6	US-10-846-172A-3
4	39	29.5	235	7	US-11-126-126-16
5	39	29.5	330	6	US-10-846-172A-7
6	39	29.5	333	6	US-10-846-172A-6
7	39	29.5	461	7	US-11-132-285-6
8	39	29.5	461	7	US-11-182-946-4
9	39	29.5	752	6	US-10-793-626-1036
10	38.5	29.2	154	6	US-10-467-657-158
11	38.5	29.2	154	6	US-10-467-657-6520
12	38.5	29.2	154	6	US-10-467-657-7634
13	37	28.0	160	6	US-10-467-657-5090
14	37	28.0	406	6	US-10-821-234-1026
15	37	28.0	425	6	US-10-793-626-110
16	37	28.0	425	6	US-10-821-234-1260
17	37	28.0	519	7	US-11-099-691-10
18	37	28.0	594	6	US-10-467-657-6376
19	36.5	27.7	658	6	US-10-467-657-4782
20	36.5	27.7	733	6	US-10-467-657-5858
21	36.5	27.7	915	6	US-10-131-826A-294
22	36.5	27.7	956	7	US-11-113-424-39
23	36	27.3	469	6	US-10-510-386-14
24	36	27.3	491	6	US-10-131-826A-278
25	36	27.3	697	7	US-11-074-176-214

Sequence 51, Appl  
Sequence 18, Appl  
Sequence 8624, Ap  
Sequence 232, App  
Sequence 180, App  
Sequence 228, App  
Sequence 1284, Ap  
Sequence 26, Appl  
Sequence 18, Appl  
Sequence 32, Appl  
Sequence 1387, Ap  
Sequence 2040, Ap  
Sequence 2, Appli  
Sequence 90, Appl  
Sequence 150, App  
Sequence 164, App  
Sequence 362, App  
Sequence 1317, Ap  
Sequence 1643, Ap  
Sequence 24, Appl  
Sequence 22, Appl  
Sequence 73, Appl  
Sequence 82, Appl  
Sequence 1155, Ap  
Sequence 650, App  
Sequence 1716, Ap  
Sequence 7, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 53, Appl  
Sequence 57, Appl  
Sequence 124, App  
Sequence 92, Appl  
Sequence 2380, Ap  
Sequence 21, Appl  
Sequence 2, Appli  
Sequence 128, App  
Sequence 7682, Ap  
Sequence 1500, Ap  
Sequence 4, Appli  
Sequence 1283, Ap  
Sequence 9149, Ap  
Sequence 378, App  
Sequence 4586, Ap  
Sequence 2624, Ap  
Sequence 522, App  
Sequence 2, Appli  
Sequence 2482, Ap  
Sequence 1070, Ap  
Sequence 7, Appli  
Sequence 1432, Ap  
Sequence 5, Appli  
Sequence 1744, Ap  
Sequence 22, Appl  
Sequence 357, App  
Sequence 9, Appli  
Sequence 2886, Ap  
Sequence 3084, Ap  
Sequence 52, Appl  
Sequence 208, App  
Sequence 3562, Ap  
Sequence 6462, Ap  
Sequence 8396, Ap  
Sequence 4, Appli  
Sequence 8, Appli  
Sequence 24, Appl  
Sequence 79, Appl  
Sequence 2140, Ap  
Sequence 13, Appl  
Sequence 2576, Ap  
Sequence 1375, Ap

Sequence 1549, Ap  
Sequence 3816, Ap

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQR 15  
| | | | | | | | | | | | | | |  
Db 2 PLHNQEVTAEDGTQR 16

ALIGNMENTS

RESULT 1

US-11-022-562-213  
; Sequence 213, Application US/11022562  
; Publication No. US20050249742A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruprecht, Ruth M.  
; APPLICANT: Shisong, Jiang  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING  
; TITLE OF INVENTION: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE  
; FILE REFERENCE: DFN-043CN  
; CURRENT APPLICATION NUMBER: US/11/022,562  
; CURRENT FILING DATE: 2004-12-22  
; PRIOR APPLICATION NUMBER: PCT/US03/20322  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 60/392718  
; PRIOR FILING DATE: 2002-06-27  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 213  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-022-562-213

Query Match 97.0%; Score 128; DB 7; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 8.8e-13;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEKCKPC 24  
| | | | | | | | | | | | | | |  
Db 316 PLHNQEVTAEDGTQRCCKCKPC 339

RESULT 2

US-11-060-646-8  
; Sequence 8, Application US/11060646  
; Publication No. US2005025555A1  
; GENERAL INFORMATION:  
; APPLICANT: Johns, Terrance Grant  
; APPLICANT: Scott, Andrew Mark  
; APPLICANT: Burgess, Anthony Wilks  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Adams, Timothy E.  
; APPLICANT: Wittup, K. Dane  
; APPLICANT: Chao, Ginger  
; TITLE OF INVENTION: EGF Receptor Epitope Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 2332-1-011N  
; CURRENT APPLICATION NUMBER: US/11/060,646  
; CURRENT FILING DATE: 2005-02-17  
; PRIOR APPLICATION NUMBER: 60/546,602  
; PRIOR FILING DATE: 2004-02-20  
; PRIOR APPLICATION NUMBER: 60/584,623  
; PRIOR FILING DATE: 2004-07-01  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: epitope peptide  
US-11-060-646-8

Query Match 60.6%; Score 80; DB 7; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;

RESULT 3

US-10-846-172A-3  
; Sequence 3, Application US/10846172A  
; Publication No. US20050257285A1  
; GENERAL INFORMATION:  
; APPLICANT: Los Alamos National Laboratory  
; APPLICANT: Gupta, Goutam  
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Pierce's Disease  
; FILE REFERENCE: S-102.340  
; CURRENT APPLICATION NUMBER: US/10/846,172A  
; CURRENT FILING DATE: 2004-05-14  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Spinachia oleracea  
US-10-846-172A-3

Query Match 29.5%; Score 39; DB 6; Length 52;  
Best Local Similarity 58.3%; Pred. No. 3;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 12 GTQRAEKCKPC 23  
| | | | | | | | | | | | | | |  
Db 41 GRRRCMCKPC 52

RESULT 4

US-11-126-126-16  
; Sequence 16, Application US/11126126  
; Publication No. US20050250696A1  
; GENERAL INFORMATION:  
; APPLICANT: Fisher F., Eric  
; APPLICANT: Edwards K., Carl  
; APPLICANT: Kieft L., Gary  
; TITLE OF INVENTION: Truncated Soluble Tumor Necrosis Factor Type-I and  
; TITLE OF INVENTION: Type-II Receptors  
; FILE REFERENCE: 02-006-A  
; CURRENT APPLICATION NUMBER: US/11/126,126  
; CURRENT FILING DATE: 2005-05-10  
; PRIOR APPLICATION NUMBER: 09/882,735  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 09/214,613  
; PRIOR FILING DATE: 1997-07-09  
; PRIOR APPLICATION NUMBER: PCT/US97/12244  
; PRIOR FILING DATE: 1997-07-09  
; PRIOR APPLICATION NUMBER: 60/039,792  
; PRIOR FILING DATE: 1997-03-04  
; PRIOR APPLICATION NUMBER: 60/039,314  
; PRIOR FILING DATE: 1997-02-07  
; PRIOR APPLICATION NUMBER: 60/037,737  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: 60/032,534  
; PRIOR FILING DATE: 1996-12-06  
; PRIOR APPLICATION NUMBER: 60/021,443  
; PRIOR FILING DATE: 1996-07-09  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-126-126-16

Query Match 29.5%; Score 39; DB 7; Length 235;

```
Best Local Similarity 43.8%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 AEDGTQRAEKCSPCA 24
Db 128 ARPGTETSDVVCCKPCA 143

RESULT 5
US-10-846-172A-7
; Sequence 7, Application US/10846172A
; Publication No. US20050257285A1
; GENERAL INFORMATION:
; APPLICANT: Los Alamos National Laboratory
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Pierce's Disease
; FILE REFERENCE: S-102,340
; CURRENT APPLICATION NUMBER: US/10/846,172A
; CURRENT FILING DATE: 2004-05-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric protein
US-10-846-172A-7

Query Match 29.5%; Score 39; DB 6; Length 330;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 12 GTQRAEKCSPC 23
Db 319 GIRRRCMSKPC 330

RESULT 6
US-10-846-172A-6
; Sequence 6, Application US/10846172A
; Publication No. US20050257285A1
; GENERAL INFORMATION:
; APPLICANT: Los Alamos National Laboratory
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Pierce's Disease
; FILE REFERENCE: S-102,340
; CURRENT APPLICATION NUMBER: US/10/846,172A
; CURRENT FILING DATE: 2004-05-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric protein
US-10-846-172A-6

Query Match 29.5%; Score 39; DB 6; Length 333;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 12 GTQRAEKCSPC 23
Db 322 GIRRRCMSKPC 333

RESULT 7
US-11-132-285-6
; Sequence 6, Application US/11132285
; Publication No. US20050244876A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR13 and TR14
; FILE REFERENCE: PF511P1
; CURRENT APPLICATION NUMBER: US/11/132,285
; CURRENT FILING DATE: 2005-05-19
; PRIOR APPLICATION NUMBER: US/10/046,433
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/261,960
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/618,570
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/144,087
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 60/149,450
; PRIOR FILING DATE: 1999-07-18
; PRIOR APPLICATION NUMBER: 60/149,712
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/153,089
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-132-285-6

Query Match 29.5%; Score 39; DB 7; Length 461;
Best Local Similarity 43.8%; Pred. No. 33;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 AEDGTQRAEKCSPCA 24
Db 150 ARPGTETSDVVCCKPCA 165

RESULT 8
US-11-182-946-4
; Sequence 4, Application US/11182946
; Publication No. US20050255100A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-4

Query Match 29.5%; Score 39; DB 7; Length 461;
Best Local Similarity 43.8%; Pred. No. 33;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 AEDGTQRAEKCSPCA 24
Db 150 ARPGTETSDVVCCKPCA 165

RESULT 9
US-10-793-626-1036
; Sequence 1036, Application US/10793626
```

```
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1036
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1036
```

```
Query Match          29.5%; Score 39; DB 6; Length 752;
Best Local Similarity 53.8%; Pred. No. 57;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 QVTAEDGTQRAE 17
Db      709 KHVTAEDGKEKAK 721
```

```
RESULT 10
US-10-467-657-158
; Sequence 158, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 158
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-158
```

```
Query Match          29.2%; Score 38.5; DB 6; Length 154;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
```

```
QY      6 EVTAEDGTQRAEKCSK 21
Db      138 KLVSEDTGTSR-EACTK 152
```

```
RESULT 11
US-10-467-657-6520
; Sequence 6520, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
```

```
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6520
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6520
```

```
Query Match          29.2%; Score 38.5; DB 6; Length 154;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
```

```
QY      6 EVTAEDGTQRAEKCSK 21
Db      138 KLVSEDTGTSR-EACTK 152
```

```
RESULT 12
US-10-467-657-7634
; Sequence 7634, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7634
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7634
```

```
Query Match          29.2%; Score 38.5; DB 6; Length 154;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
```

```
QY      6 EVTAEDGTQRAEKCSK 21
Db      138 KLVSEDTGTSR-EACTK 152
```

```
RESULT 13
US-10-467-657-5090
; Sequence 5090, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
```

; SEQ ID NO 5090  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-5090

Query Match 28.0%; Score 37; DB 6; Length 160;  
Best Local Similarity 50.0%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 0; Indels 5; Gaps 0;

Qy 13 TQRAEKSKPCA 24  
|||:|  
Db 134 THRRTCTAPCA 145

## RESULT 14

US-10-821-234-1026  
; Sequence 1026, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 1026  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-821-234-1026

Query Match 28.0%; Score 37; DB 6; Length 406;  
Best Local Similarity 36.4%; Pred. No. 60;  
Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAED 11  
|:|:|:|:|  
Db 31 PINNEDITYDE 41

## RESULT 15

US-10-793-626-110  
; Sequence 110, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 110  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-10-793-626-110

Query Match 28.0%; Score 37; DB 6; Length 425;  
Best Local Similarity 42.1%; Pred. No. 63;  
Matches 8; Conservative 3; Mismatches 0; Indels 8; Gaps 0;

Qy 2 LHNQEVTAEDGTORAEKCS 20  
|||:|:|:|  
Db 230 LHRGERSGEEVAQRMERAS 248

## RESULT 16

US-10-821-234-1260  
; Sequence 1260, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 1260  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-821-234-1260

Query Match 28.0%; Score 37; DB 6; Length 425;  
Best Local Similarity 46.2%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 HNOEVTAEDGTQR 15  
||:|:|  
Db 327 HNETILASSGTR 339

## RESULT 17

US-11-099-691-10  
; Sequence 10, Application US/11099691  
; Publication No. US20050260644A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: YANG, Junming  
; TITLE OF INVENTION: CELL SIGNALING PROTEINS  
; FILE REFERENCE: PF-0521 PCT  
; CURRENT APPLICATION NUMBER: US/11/099,691  
; CURRENT FILING DATE: 2005-04-06  
; PRIOR APPLICATION NUMBER: US/09/700,444  
; PRIOR FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: 60/085,343  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/098,010  
; PRIOR FILING DATE: 1998-08-26  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PERL Program  
; SEQ ID NO 10  
; LENGTH: 519  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte Clone 1652936  
US-11-099-691-10

Query Match 28.0%; Score 37; DB 7; Length 519;  
Best Local Similarity 54.5%; Pred. No. 79;  
Matches 6; Conservative 4; Mismatches 1; Indels 1; Gaps 0;

QY 1 PLHNQEVTAED 11  
|:|:|:|:|:  
DB 273 PVENQDVTSD 283

RESULT 18  
US-10-467-657-6376  
; Sequence 6376, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 6376  
; LENGTH: 594  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6376

Query Match 28.0%; Score 37; DB 6; Length 594;  
Best Local Similarity 41.2%; Pred. No. 91;  
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 HNOEVTAE DGTQRAEKC 19  
|:|:|:|:|:  
DB 549 HHQPNQDDGEYAEKC 565

RESULT 19  
US-10-467-657-4782  
; Sequence 4782, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 4782  
; LENGTH: 658  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-4782

Query Match 27.7%; Score 36.5; DB 6; Length 658;  
Best Local Similarity 45.5%; Pred. No. 1.2e+02;  
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

QY 3 HNOEVTAE DGTQRAEKCSPCA 24  
|:|:|:|:|:  
DB 120 HATQVKAEDG-----KLSSPTA 136

RESULT 20  
US-10-467-657-5858  
; Sequence 5858, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 5858  
; LENGTH: 733  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-5858

Query Match 27.7%; Score 36.5; DB 6; Length 733;  
Best Local Similarity 56.2%; Pred. No. 1.4e+02;  
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 4 NOEVTVA-EDGTQRAEK 18  
|:|:|:|:|:  
DB 485 NQDLTAVGHGTQVEE 500

RESULT 21  
US-10-131-826A-294  
; Sequence 294, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C128  
; CURRENT APPLICATION NUMBER: US/10/131,826A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122

```
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 294
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-294

Query Match      27.7%; Score 36.5; DB 6; Length 915;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches      8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Oy      7 VTAEDGTQRAEKCSK 21
Db      597 VLAEDG-RRCKKCTE 610

RESULT 22
US-11-113-424-39
; Sequence 39, Application US/11113424
; Publication No. US20050260713A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/11/113,424
; CURRENT FILING DATE: 2005-04-21
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-424-39

Query Match      27.7%; Score 36.5; DB 7; Length 956;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches      8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Oy      7 VTAEDGTQRAEKCSK 21
Db      638 VLAEDG-RRCKKCTE 651

RESULT 23
US-10-510-386-14
```

```
; Sequence 14, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294,204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-14

Query Match      27.3%; Score 36; DB 6; Length 469;
Best Local Similarity 41.2%; Pred. No. 1e+02;
Matches      7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy      6 EVTAEDGTQRAEKCSKP 22
Db      232 EITADIFTKKDEETGKP 248

RESULT 24
US-10-131-826A-278
; Sequence 278, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
```

```
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 278
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-278

Query Match      27.3%; Score 36; DB 6; Length 491;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 AEDGTQRAEKS 20
Db 39 ATDGKEAKKA 50

RESULT 25
US-11-074-176-214
; Sequence 214, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 214
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-214

Query Match      27.3%; Score 36; DB 7; Length 697;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 QEVTAEADGTQRAE 17
Db 430 QKLTEEDPTFRAE 442

RESULT 26
US-11-137-465-51
; Sequence 51, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
```

```
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-137-465-51

Query Match      27.3%; Score 36; DB 7; Length 753;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 NOEVTAEADGTQRAEKS 20
Db 698 NFEVAESDFTNNAMKCN 714

RESULT 27
US-11-078-735-18
; Sequence 18, Application US/11078735
; Publication No. US20050261477A1
; GENERAL INFORMATION:
; APPLICANT: CHAMPION, BRIAN ROBERT
; APPLICANT: LENNARD, ANDREW CHRISTOPHER
; APPLICANT: MCKENZIE, GRAHAME JAMES
; APPLICANT: TUGAL, TAMARA
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND MEDICAL TREATMENTS
; TITLE OF INVENTION: COMPRISING NOTCH LIGAND PROTEINS
; FILE REFERENCE: 674525-2019
; CURRENT APPLICATION NUMBER: US/11/078,735
; CURRENT FILING DATE: 2005-03-10
; PRIOR APPLICATION NUMBER: PCT/GB03/03908
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: PCT/GB03/03285
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: PCT/GB03/01525
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: GB 0300234.2
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: PCT/GB02/05137
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/GB02/05133
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: GB 0220912.0
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: GB 0220913.8
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 18
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-078-735-18

Query Match      26.9%; Score 35.5; DB 7; Length 618;
Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 11 DGTQRAEKC-SKPCA 24
Db 425 DCRERADPCAARPCA 439

RESULT 28
US-10-467-657-8624
; Sequence 8624, Application US/10467657
```



; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 8624  
; LENGTH: 26  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-8624

Query Match 26.5%; Score 35; DB 6; Length 26;  
Best Local Similarity 50.0%; Pred. No. 6;  
Matches 7; Conservative 1; Mismatches 6; Indels 6; Gaps 0;

Qy 10 EDGTQRAEKCSPC 23  
Db 1 EDGEDADGRGSKPC 14

RESULT 29  
; US-10-793-626-252  
; Sequence 252, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 252  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-10-793-626-252

Query Match 26.5%; Score 35; DB 6; Length 164;  
Best Local Similarity 66.7%; Pred. No. 46;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 NQEVTAEDG 12  
Db 68 NEESAEDG 76

RESULT 30  
US-10-485-517-180  
; Sequence 180, Application US/10485517  
; Publication No. US20050256299A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Biosynexus Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides  
; FILE REFERENCE: P100629W0

; CURRENT APPLICATION NUMBER: US/10/485,517  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: GB 0118825.9  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: GB 0200349.9  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 424  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 180  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-485-517-180

Query Match 26.5%; Score 35; DB 6; Length 216;  
Best Local Similarity 32.1%; Pred. No. 62;  
Matches 9; Conservative 8; Mismatches 7; Indels 4; Gaps 2;

Qy 1 PLHNQEV-TAED--GTOAEKCKSPCA 24  
Db 147 PLNSVDINTKEDFKATIERSDSCAVPAA 174

RESULT 31  
US-10-510-386-228  
; Sequence 228, Application US/10510386  
; Publication No. US20050244922A1  
; GENERAL INFORMATION:  
; APPLICANT: Andersen, Jens Tonne  
; APPLICANT: Clausen, Ib Groth  
; APPLICANT: Jorgensen, Steen Troels  
; APPLICANT: Olsen, Peter Bjarke  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Improved Bacillus Host Cell  
; FILE REFERENCE: 10294.204-US  
; CURRENT APPLICATION NUMBER: US/10/510,386  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 248  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 228  
; LENGTH: 229  
; TYPE: PRT  
; ORGANISM: Bacillus licheniformis  
US-10-510-386-228

Query Match 26.5%; Score 35; DB 6; Length 229;  
Best Local Similarity 43.8%; Pred. No. 66;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 EVTAEDGTQAEKCKSK 21  
Db 66 ELKKDAGDQKAEKTAK 81

RESULT 32  
US-10-821-234-1284  
; Sequence 1284, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 1284  
; LENGTH: 274

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-821-234-1284

Query Match 26.5%; Score 35; DB 6; Length 274;  
Best Local Similarity 33.3%; Pred. No. 80;  
Matches 6; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 4 NQEVTAEDGTQRAEKCSK 21

Db 48 SQEVTREDESTRSEEVTR 65

RESULT 33

US-10-131-826A-26  
; Sequence 26, Application US/10131826A  
; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P330R1C128

; CURRENT APPLICATION NUMBER: US/10/131,826A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059588

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 26

; LENGTH: 296

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-131-826A-26

Query Match 26.5%; Score 35; DB 6; Length 296;  
Best Local Similarity 46.2%; Pred. No. 88;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 HNOEVTAEEDGTQR 15

Db 116 HNYETTAKESLRR 128

RESULT 34

US-11-102-240-18

; Sequence 18, Application US/11102240

; Publication No. US20050260647A1

; GENERAL INFORMATION:

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRES

; TITLE OF INVENTION: ESOPHAGEAL TUMOR

; FILE REFERENCE: P3230R1C106C

; CURRENT APPLICATION NUMBER: US/11/102,240

; CURRENT FILING DATE: 2005-04-08

; PRIOR APPLICATION NUMBER: 10/063662

; PRIOR FILING DATE: 2002-05-07

; PRIOR APPLICATION NUMBER: 10/006867

; PRIOR FILING DATE: 2001-12-06

; PRIOR APPLICATION NUMBER: PCT/US00/23328

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 60/170262

; PRIOR FILING DATE: 199-12-09

; NUMBER OF SEQ ID NOS: 170

; SEQ ID NO 18

; LENGTH: 301

; TYPE: PRT

; ORGANISM: Homo Sapien

US-11-102-240-18

Query Match 26.5%; Score 35; DB 7; Length 301;

Best Local Similarity 43.8%; Pred. No. 89;

Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 VTAEDGTQRAEKCSKP 22

Db 115 LTAIEGTAHGEPCHFP 130

RESULT 35

US-11-074-176-32

; Sequence 32, Application US/11074176

; Publication No. US20050250135A1

; GENERAL INFORMATION:

; APPLICANT: Klaenhammer, Todd R.

; APPLICANT: Russell, William M.

; APPLICANT: Altermann, Eric

; APPLICANT: McAuliffe, Olivia

; APPLICANT: Perill, Andrea Azcarate

; TITLE OF INVENTION: Nucleic Acid Sequences Encoding

; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore

; FILE REFERENCE: 5051-694

; CURRENT APPLICATION NUMBER: US/11/074,176

; CURRENT FILING DATE: 2005-03-07

; PRIOR APPLICATION NUMBER: 60/551,161

; PRIOR FILING DATE: 2004-03-08

; NUMBER OF SEQ ID NOS: 381

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 32

; LENGTH: 333

; TYPE: PRT

; ORGANISM: Lactobacillus acidophilus

US-11-074-176-32

Query Match 26.5%; Score 35; DB 7; Length 333;

Best Local Similarity 43.8%; Pred. No. 1e+02;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LHNQEVTAEDGTQRAE 17



```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-90

Query Match      26.5%; Score 35; DB 6; Length 468;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LHNQEVTA 10
Db      31 LHNRELSAE 39

RESULT 40
US-11-102-240-150
; Sequence 150, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P323OR1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 150
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-150

Query Match      26.5%; Score 35; DB 7; Length 468;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LHNQEVTA 10
Db      31 LHNRELSAE 39

Search completed: December 3, 2005, 08:04:41
Job time : 8.09091 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:36:13 ; Search time 25.0909 Seconds  
(without alignments)  
92.033 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLNNQEVTAEDGTQRAEKSKPCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR\_80.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	1255	1 A24571	protein-tyrosine k
2	114	86.4	1260	1 TVRTNU	protein-tyrosine k
3	113	85.6	1254	2 I48161	p-185 precursor -
4	50	37.9	686	2 A34612	zinc finger protei
5	49	37.1	1210	1 GQHUE	epidermal growth f
6	48	36.4	214	2 H90840	hypothetical prote
7	48	36.4	214	2 G85698	hypothetical prote
8	48	36.4	644	2 A36325	epidermal growth f
9	48	36.4	644	2 B71409	hypothetical prote
10	48	36.4	1210	2 A53183	epidermal growth f
11	48	36.4	1224	2 T19749	hypothetical prote
12	47	35.6	774	2 T00488	hypothetical prote
13	46	34.8	286	2 T05061	hypothetical prote
14	46	34.8	651	2 JC7705	death receptor-6 -
15	46	34.8	1390	2 T30346	insulin receptor -
16	46	34.8	2647	2 T28161	hypothetical prote
17	45	34.1	196	2 T49548	hypothetical prote
18	45	34.1	379	2 C84577	probable nucleosom
19	45	34.1	620	2 JU0273	hemagglutinin - me
20	45	34.1	742	2 E84285	hypothetical prote
21	45	34.1	873	2 H96503	protein F9C16.17 l
22	44.5	33.7	499	2 S70113	hypothetical prote
23	44	33.3	91	2 E81715	ferredoxin [2Fe-2S
24	44	33.3	268	2 AE3311	cytochrome c1 limp
25	44	33.3	288	2 PH1917	FL-160-1 protein -
26	44	33.3	577	2 T09024	proline-rich prote
27	44	33.3	609	2 T40625	single-stranded DN
28	44	33.3	717	2 T04434	replication protei
29	44	33.3	1051	2 C95367	conserved hypothet

30	44	33.3	1422	2 T42636	protein-tyrosine-p
31	43.5	33.0	606	1 W1WLB5	E1 protein - human
32	43.5	33.0	681	2 I52603	MPS1 protein - mou
33	43.5	33.0	707	2 S60588	drebrin A - rat
34	43	32.6	91	2 G71562	ferredoxin [2Fe-2S
35	43	32.6	199	1 I39734	CBS domain-contain
36	43	32.6	204	2 AC1902	hypothetical prote
37	43	32.6	296	2 T20005	hypothetical prote
38	43	32.6	450	2 JC7266	3',5'-cyclic-nucle
39	43	32.6	516	2 S28060	serum response fac
40	43	32.6	585	2 JC6556	Y chromosome sex-d
41	43	32.6	655	2 T34219	hypothetical prote
42	43	32.6	701	2 T17243	hypothetical prote
43	43	32.6	1039	2 S46347	pol polyprotein -
44	43	32.6	1101	2 T16840	hypothetical prote
45	43	32.6	1293	2 B40025	maleless protein,
46	43	32.6	2185	1 S60200	acetyl-CoA carboxy
47	43	32.6	2396	2 T13714	kakapo gene protei
48	43	32.6	3343	2 S44887	ZK112.7 protein -
49	43	32.6	5138	2 B96695	hypothetical prote
50	42.5	32.2	275	2 S21348	probable pol polyp
51	42	31.8	77	1 XKTO	metallocarboxypept
52	42	31.8	163	2 A97609	crea protein limpo
53	42	31.8	163	2 AE2831	conserved hypothet
54	42	31.8	215	2 A60166	hemiferrin - bovin
55	42	31.8	216	1 A39684	hemiferrin - rat
56	42	31.8	222	2 F95367	hypothetical prote
57	42	31.8	284	2 T40927	hypothetical wflil
58	42	31.8	297	2 S51278	alternative respir
59	42	31.8	353	2 T04094	ubiquinol-cytochro
60	42	31.8	450	1 C29413	prolyl-tRNA synthe
61	42	31.8	499	2 D75416	uroconase (urocana
62	42	31.8	559	2 G83897	hypothetical prote
63	42	31.8	603	2 S76615	hypothetical prote
64	42	31.8	687	2 T30958	hypothetical prote
65	42	31.8	1082	2 T05075	hypothetical prote
66	42	31.8	1125	2 B41206	microtubule-associ
67	42	31.8	1941	2 T23979	hypothetical prote
68	42	31.8	1943	2 T23986	hypothetical prote
69	41.5	31.4	273	2 T01985	zinc-finger protei
70	41.5	31.4	328	2 AH1226	N-acetyluramoyl-L
71	41	31.1	194	2 JS0664	interferon precurs
72	41	31.1	201	2 S44540	DNA-directed DNA p
73	41	31.1	240	2 H81420	probable iron-bind
74	41	31.1	365	2 T25917	hypothetical prote
75	41	31.1	453	2 S59401	hypothetical prote
76	41	31.1	520	2 T34383	hypothetical prote
77	41	31.1	550	2 A11522	flagellar basal-bo
78	41	31.1	594	2 S74576	probable flavoprot
79	41	31.1	617	2 D86248	protein T23J18.15
80	41	31.1	704	2 A34337	propionyl-CoA carb
81	41	31.1	843	2 S44868	kinesin heavy chal
82	41	31.1	958	2 T20621	hypothetical prote
83	41	31.1	975	1 A31497	hypothetical prote
84	41	31.1	1002	2 T09438	toxR-activated lip
85	41	31.1	1013	2 B82276	ToxR-activated gen
86	41	31.1	1030	2 T18374	B-cell receptor pr
87	41	31.1	1172	2 F96503	protein F9C16.13 l
88	41	31.1	1536	1 R85Y53	regulatory protein
89	41	31.1	1599	2 T15854	hypothetical prote
90	41	31.1	1984	2 T13171	probable vitellogen
91	40.5	30.7	171	2 T72578	hypothetical prote
92	40.5	30.7	224	2 T06736	GTP-binding protei
93	40.5	30.7	403	1 G70311	histidine-tRNA lig
94	40.5	30.7	621	1 S20145	replication factor
95	40	30.3	129	2 S44751	C0664.4 protein -
96	40	30.3	144	2 A39499	6-pyruvoyltetrahyd
97	40	30.3	156	2 I40704	gene eae ORF homo
98	40	30.3	156	2 D86045	hypothetical prote
99	40	30.3	156	2 H91198	Cesr protein limpo
100	40	30.3	156	2 I41196	Tir chaperone - Es

## ALIGNMENTS

## RESULT 1

A24571  
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human  
N;Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB  
C;Species: Homo sapiens (man)  
C;Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 05-Oct-2004  
C;Accession: A24571; A25491; A44188; B44188; 159509; I57622  
R;Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;  
Nature 319, 230-234, 1986  
A;Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor  
A;Reference number: A24571; MUID:86118663; PMID:3003577  
A;Accession: A24571  
A;Molecule type: mRNA  
A;Residues: 1-1255 <YAM>  
A;Cross-references: UNIPROT:P04626; UNIPARC:UPI000003F55F; GB:X03363; NID:g31197; PIDN:C  
R;Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A;Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal  
A;Reference number: A25491; MUID:86016729; PMID:2995967  
A;Accession: A25491  
A;Molecule type: DNA  
A;Residues: 737-1031 <SEM>  
A;Cross-references: UNIPARC:UPI000016A8A7; GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:  
R;Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.  
Science 230, 1132-1139, 1985  
A;Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromo  
A;Reference number: A44188; MUID:86070181; PMID:2999974  
A;Accession: A44188  
A;Molecule type: DNA  
A;Residues: 740-910 <COUL>  
A;Cross-references: UNIPARC:UPI000016AA26; GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:  
A;Accession: B44188  
A;Molecule type: mRNA  
A;Residues: 1-517, 'RALL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A;Cross-references: UNIPARC:UPI00001725C7; GB:M11730; NID:g183986  
R;King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A;Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A;Reference number: I59509; MUID:85272597; PMID:2992089  
A;Accession: I59509  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 832-909 <REX>  
A;Cross-references: UNIPARC:UPI0000070A3F; GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:  
R;Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A;Title: Human HSR2 (neu) promoter: evidence for multiple mechanisms for transcriptional  
A;Reference number: I57622; MUID:87286898; PMID:3039351  
A;Accession: I57622  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-191 <TAL>  
A;Cross-references: UNIPARC:UPI0000000427; GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:  
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
C;Genetics:  
A;Gene: GDB:ERBB2; NCL; NEU; HER-2  
A;Cross-references: GDB:I20613; OMIM:164870  
A;Map position: 17q21.1-17q21.1  
A;Introns: 25/1; 75/3; 147/1; 883/3  
A;Note: the list of introns is incomplete  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
inase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>  
F;22-653/Domain: extracellular #status predicted <EXT>  
F;70-304/Domain: EGF receptor extracellular domain repeat <BE1>  
F;395-605/Domain: EGF receptor extracellular domain repeat <EE2>

## ALIGNMENTS

F;654-675/Domain: transmembrane #status predicted <TMM>  
F;676-1255/Domain: intracellular #status predicted <INT>  
F;718-983/Domain: protein kinase homology <KIN>  
F;726-734/Region: protein kinase ATP-binding motif  
F;68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predic  
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
F;753/Active site: Lys #status predicted  
F;1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)  
Query Match 97.0%; Score 128; DB 1; Length 1255;  
Best Local Similarity 95.8%; Pred. No. 6.7e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
|||||  
DB 316 PLHNQEVTAEDGTQRCCKCKPCA 339  
RESULT 2  
TVRTNU  
protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 05-Oct-2004  
C;Accession: A24562; A61204  
R;Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.  
Nature 319, 226-230, 1986  
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.  
A;Reference number: A24562; MUID:86118662; PMID:3945311  
A;Accession: A24562  
A;Molecule type: mRNA  
A;Residues: 1-1260 <BAR>  
A;Cross-references: UNIPROT:P06494; UNIPARC:UPI0000161B83; EMBL:X03362; NID:g56745; PID:  
R;Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,  
Carcinogenesis 12, 1975-1978, 1991  
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no  
2-thiazolylformamide or N-methyl-N-nitrosourea.  
A;Reference number: A61204; MUID:92035293; PMID:1682063  
A;Accession: A61204  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 637-663, 'V', 665-702 <MAS>  
A;Cross-references: UNIPARC:UPI00001725C8  
A;Note: authors translated the codon GCA for residue 25 as Val  
C;Genetics:  
A;Gene: neu  
C;Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph  
F;1-13/Domain: signal sequence #status predicted <SIG>  
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>  
F;658-680/Domain: transmembrane #status predicted <TMN>  
F;723-988/Domain: protein kinase homology <KIN>  
F;731-739/Region: protein kinase ATP-binding motif  
F;71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;691/Binding site: phosphate (Thr) (covalent) #status predicted  
F;758/Active site: Lys #status predicted  
F;882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted  
Query Match 86.4%; Score 114; DB 1; Length 1260;  
Best Local Similarity 87.5%; Pred. No. 8.4e-09;  
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 PLHNQEVTAEDGTQRAEKCKPCA 24  
|||||  
DB 320 PNHNQEVTAEDGTQRCCKCKPCA 343  
RESULT 3  
I48161  
p-185 precursor - golden hamster  
C;Species: Mesocricetus auratus (golden hamster)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Oct-2004  
C;Accession: I48161  
R;Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa

Gene 140, 251-255, 1994  
A;Title: Cloning and activation of the Syrian hamster neu proto-oncogene.  
A;Reference number: 148161; MUID:94193007; PMID:7908275  
A;Accession: 148161  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1254 <RES>  
A;Cross-references: UNIPROT:Q60553; UNIPARC:UPI000012A111; GB:D16295; NID:g493236; PIDN:  
C;Genetics:  
A;Gene: neu  
C;Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology  
C;Keywords: ATP  
F:718-983/Domain: protein kinase homology <KIN>  
F:726-734/Region: protein kinase ATP-binding motif

Query Match 85.6%; Score 113; DB 2; Length 1254;  
Best Local Similarity 87.5%; Pred. No. 1.2e-08;  
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PLNNQVTAEDGTQRAEKCCKP 24  
Db 316 PLNNQVTAEDGTQRAEKCCKP 339

RESULT 4  
A34612  
N;Alternate names: zinc finger protein ZNF7 - human  
C;Species: Homo sapiens (man)  
C;Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 09-Jul-2004  
R;Lania, L.; Donti, E.; Pannuti, A.; Pascucci, A.; Pengue, G.; Feliciello, I.; La Mantia  
Genomic 6, 333-340, 1990  
A;Title: cDNA isolation, expression analysis, and chromosomal localization of two human  
A;Reference number: A34612; MUID:90169993; PMID:2106481  
A;Accession: A34612  
A;Molecule type: mRNA  
A;Residues: 1-686 <LAN>  
A;Cross-references: UNIPROT:P17097; UNIPARC:UPI000013C3F8; GB:M29580; NID:g340445; PIDN:  
R;Bray, P.; Lichter, P.; Thiesen, H.J.; Ward, D.C.; David, I.B.  
Proc. Natl. Acad. Sci. U.S.A. 88, 9563-9567, 1991  
A;Title: Characterization and mapping of human genes encoding zinc finger proteins.  
A;Reference number: A56409; MUID:92052132; PMID:1946370  
A;Accession: A56409  
A;Molecule type: DNA  
A;Residues: 425-589 <BRA>  
A;Cross-references: UNIPARC:UPI0000178A54; GB:M77170  
R;Thiesen, H.J.  
submitted to the EMBL Data Library, March 1990  
A;Reference number: S10397  
A;Accession: S10421  
A;Molecule type: mRNA  
A;Residues: 413-468 <THI>  
A;Cross-references: UNIPARC:UPI000016ABC5; EMBL:X52335; NID:g34165; PIDN:CAA36561.1; PIDN:  
R;Thiesen, H.J.  
New Biol. 2, 363-374, 1990  
A;Title: Multiple genes encoding zinc finger domains are expressed in human T cells.  
A;Reference number: 137949; MUID:91145339; PMID:2288909  
A;Accession: 137972  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 413-468 <RES>  
A;Cross-references: UNIPARC:UPI000016ABC5; EMBL:X52335; NID:g34165; PIDN:CAA36561.1; PIDN:  
C;Genetics:  
A;Gene: GDB:ZNF7  
A;Cross-references: GDB:120509; OMIM:194531  
A;Map position: 8q24.3-8q24.3  
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology  
C;Keywords: DNA binding, zinc finger

Query Match 37.9%; Score 50; DB 2; Length 686;  
Best Local Similarity 47.6%; Pred. No. 18;  
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PLNNQVTAEDGTQRAEKCCK 21  
Db 183 PLESQGESAGMSQRCCECK 203

RESULT 5  
QOHUE  
epidermal growth factor receptor precursor - human  
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
C;Species: Homo sapiens (man)  
C;Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 05-Oct-2004  
C;Accession: A00641; A25772; A38672; A00642; A43615; A23062; A05281; A60143; A3  
R;Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.;  
rg, P.H.  
Nature 309, 418-425, 1984  
A;Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression o  
A;Reference number: A00641; MUID:84219729; PMID:6328312  
A;Accession: A00641  
A;Molecule type: mRNA  
A;Residues: 1-1210 <ULL>  
A;Cross-references: UNIPROT:P00533; UNIPARC:UPI0000050F30; EMBL:X00588; NID:g31113; PID  
A;Note: the authors translated the codon AAG for residue 540 as Asn  
R;Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985  
A;Title: Characterization and sequence of the promoter region of the human epidermal gr  
A;Reference number: A25772; MUID:85270438; PMID:2991899  
A;Accession: A25772  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-29 <IGH>  
A;Cross-references: UNIPARC:UPI000016A882; GB:M11234; NID:g181981; PIDN:AAA52370.1; PID  
R;Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.  
Oncogene Res. 1, 375-396, 1987  
A;Title: The human EGF receptor gene: structure of the 110 kb locus and identification  
A;Reference number: S30024; MUID:88217333; PMID:3329716  
A;Accession: S30024  
A;Molecule type: DNA  
A;Residues: 1-29 <HA2>  
A;Cross-references: UNIPARC:UPI000016A882; EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PI  
R;Haley, J.D.; Waterfield, M.D.  
J. Biol. Chem. 266, 1746-1753, 1991  
A;Title: Contributory effects of de Novo transcription and premature transcript termina  
A;Reference number: A38672; MUID:91107677; PMID:1988448  
A;Accession: A38672  
A;Molecule type: DNA  
A;Residues: 1-29 <HAL>  
A;Cross-references: UNIPARC:UPI000016A882; GB:M38425; NID:g181977; PIDN:AAA63171.1; PID  
A;Experimental source: carcinoma cell line A431-7  
R;Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Me  
Nature 309, 806-810, 1984  
A;Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs  
A;Reference number: A00642; MUID:84245835; PMID:6330563  
A;Accession: A00642  
A;Molecule type: mRNA  
A;Residues: 'RCAWRA' 150-187, 'KSVIOAV' 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-32  
' 798-799, 'TD', 802-811, 'R', 813-942 <XUY>  
A;Cross-references: UNIPARC:UPI00001725BD  
A;Experimental source: A431 human carcinoma cells, which have large numbers of EGF rece  
R;Lin, C.R.; Chen, W.S.; Krueger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma, I  
Science 224, 843-848, 1984  
A;Title: Expression cloning of human EGF receptor complementary DNA: gene amplification  
A;Reference number: A43615; MUID:84196372; PMID:6326261  
A;Accession: A43615  
A;Molecule type: mRNA  
A;Residues: 713-964 <LIN>  
A;Cross-references: UNIPARC:UPI00001725BE  
A;Experimental source: epidermoid carcinoma cell line A431  
R;Simmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
A;Reference number: A23062; MUID:85046483; PMID:6093780  
A;Accession: A23062  
A;Molecule type: mRNA







A;Description: A novel protein from carrot embryo.

A;Reference number: T14284

A;Accession: T00488

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: mRNA

A;Residues: 1-774 <XUZ>

A;Cross-references: UNIPROT:O80379; UNIPARC:UPI00000A89BA; EMBL:AB012704; NID:G3551248;

A;Experimental source: embryogenic callus

C;Genetics:

A;Note: 184-1d10

Query Match 35.6%; Score 47; DB 2; Length 774;

Best Local Similarity 40.0%; Pred. No. 56;

Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 4 NORVTAEDGTQRAEKCKPC 23

||:|:|:|:|:|:|:|:|

Db 313 NBEITPQDESTRAKSPKLC 332

RESULT 13

T05061

hypothetical protein M3E9.140 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T05061

R;Bevan, M.; Vandenbol, M.; Jallet, C.; Portetelle, D.; Hoheisel, J.; Mewes, H.W.; Mayer

submitted to the Protein Sequence Database, March 1999

A;Reference number: Z15396

A;Accession: T05061

A;Molecule type: DNA

A;Residues: 1 286 <BEV>

A;Cross-references: UNIPROT:O65591; UNIPARC:UPI00000A63D0; EMBL:AL022223

A;Experimental source: cultivar Columbia; BAC clone M3E9

C;Genetics:

A;Map position: 4

A;Introns: 100/1; 129/3; 161/1; 181/3; 208/1; 241/3; 260/2

A;Note: M3E9.140

Query Match 34.8%; Score 46; DB 2; Length 286;

Best Local Similarity 38.1%; Pred. No. 30;

Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 3 HNOEVTAEDGTQRAEKCKPC 23

||:|:|:|:|:|:|:|:|

Db 192 HVAHLKPSDGSATQCDKPC 212

RESULT 14

JC7705

death receptor-6 - chicken

C;Species: Gallus gallus (chicken)

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Jul-2004

C;Accession: JC7705

K;Bridgham, J.T.; Bohe, J.; Goetz, F.W.; Johnson, A.L.

Biochem. Biophys. Res. Commun. 284, 1109-1115, 2001

A;Title: Conservation of death receptor-6 in avian and piscine vertebrates.

A;Reference number: JC7705; MUID:21308433; PMID:11414698

A;Accession: JC7705

A;Molecule type: mRNA

A;Residues: 1-651 <BRI>

A;Cross-references: UNIPROT:Q98SM6; UNIPARC:UPI000017CBA0; GB:AF349908

C;Comment: This receptor, a member of the tumor necrosis factor receptor family, belongs

to the TNF receptor superfamily, and is involved in cell death and/or survival signaling cascade.

C;Genetics:

A;Gene: dr-6

C;Superfamily: tumor necrosis factor receptor type 1 (TNFR1); NGF receptor repeat homolo

C;Keywords: ovary

F;1-21/Domain: signal sequence #status predicted <SIG>

F;52-196/Domain: extracellular cysteine-rich, ligand-binding #status predicted <EC1>

F;332-350/Domain: transmembrane #status predicted <TM>

F;410-475/Domain: death domain #status predicted <DE0>

F;551-651/Region: conserved cytoplasmic #status predicted

Query Match 34.8%; Score 46; DB 2; Length 651;

Best Local Similarity 50.0%; Pred. No. 67;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 10 EDGTQRAEKCKPC 23

||:|:|:|:|:|:|:|:|

Db 85 ENGIERCHPCRKPC 98

RESULT 15

T30346

insulin receptor - yellow fever mosquito

C;Species: Aedes aegypti (yellow fever mosquito)

C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T30346

R;Graf, R.; Neuenschwander, S.; Brown, M.R.; Ackermann, U.

Insect Mol. Biol. 6, 151-163, 1996

A;Title: Insulin mediated secretion of ecdysteroids from mosquito ovaries and molecular

A;Reference number: Z20834

A;Accession: T30346

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: mRNA

A;Residues: 1-1390 <GRA>

A;Cross-references: UNIPROT:Q93105; UNIPARC:UPI000012D6AB; EMBL:U72939; NID:G1620749; P

C;Superfamily: insulin receptor; protein kinase homology

Query Match 34.8%; Score 46; DB 2; Length 1390;

Best Local Similarity 70.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 15 RAEKCKPC 24

||:|:|:|:|:|:|:|:|

Db 1048 RGEKCNQPCA 1057

RESULT 16

T28161

hypothetical protein FCR3-varT11-1 - malaria parasite (Plasmodium falciparum) (fragment)

C;Species: Plasmodium falciparum

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T28161

R;Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellem, T.E.; Scherf,

Mol. Cell. Biol. 17, 604-611, 1997

A;Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.

A;Reference number: Z20483; MUID:97154495; PMID:9001213

A;Accession: T28161

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-2647 <HER>

A;Cross-references: UNIPROT:P90580; UNIPARC:UPI0000083525; EMBL:U67959; NID:G1794255; P

A;Experimental source: strain FCQ27/PNG

C;Genetics:

A;Introns: 2158/3

A;Note: FCR3-varT11-1

Query Match 34.8%; Score 46; DB 2; Length 2647;

Best Local Similarity 47.6%; Pred. No. 2.6e+02;

Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 LHNQEVTAEDGTQRAEKCKSP 22

||:|:|:|:|:|:|:|:|

Db 727 LSDSEDEEDIPQRQNKCAK 747

RESULT 17

T49548

hypothetical protein B21J21.260 [imported] - Neurospora crassa

C;Species: Neurospora crassa

C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000

C;Accession: T49548

R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A:Reference number: 225022

A:Accession: T49548

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <SCH>

A:Cross-references: UNIPARC:UPI0000179D39; EMBL:AL355929; GSPDB:GN00116; NCSP:B21J21.260

A:Experimental source: BAC clone B21J21; strain OR74A

C:Genetics:

A:Gene: NCSP:B21J21.260

A:Map position: 6

C:Superfamily: Neurospora crassa hypothetical protein B21J21.260

Query Match

Best Local Similarity 34.1%; Score 45; DB 2; Length 196;

Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Oy 1 PLHNQEVTAEDGTQRAEKCS 20

Db 155 PLHQRLTSLDAGSQPKNCA 174

|||||:|:|:|:|:|:|

RESULT 18

C84577

Probable nucleosome assembly protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: C84577

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nietman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84577

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <STO>

A:Cross-references: UNIPROT:Q9ZUP3; UNIPARC:UPI00000AC55A; GB:AE002093; NID:g4191778; PI

C:Genetics:

A:Gene: At2g19480

A:Map position: 2

Query Match

Best Local Similarity 34.1%; Score 45; DB 2; Length 379;

Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 2 LHNQEVTAEDGTQRAEKCSK 21

Db 134 LKNEITAEETERDEGALK 153

|||||:|:|:|:|:|:|

RESULT 19

JU0273

hemagglutinin - measles virus

C:Species: measles virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C:Accession: JU0273

R:Komase, K.; Hago, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.

Virus Genes 4, 163-172, 1990

A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe

A:Reference number: JU0273; MUID:90385701; PMID:2402882

A:Accession: JU0273

A:Molecule type: mRNA

A:Residues: 1-620 <KOM>

A:Cross-references: UNIPROT:P28081; UNIPARC:UPI0000178694

A:Experimental source: strain Yamagata-1

C:Genetics:

A:Gene: H

C:Superfamily: measles virus hemagglutinin

C:Keywords: Glycoprotein; transmembrane protein

F:35-58/Domain: transmembrane #status predicted <TMM>

F:168,187,200,215,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 34.1%; Score 45; DB 2; Length 620;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 7 VTAEDGTQRAEKC 19

Db 608 VTREDGTNRQSC 620

|||||:|:|:|:|:|:|

RESULT 20

E84285

hypothetical protein Vng1303c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: E84285

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: E84285

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-742 <STO>

A:Cross-references: UNIPROT:Q9HQ68; UNIPARC:UPI000006385E; GB:AE004437; NID:gl0580823;

C:Genetics:

A:Gene: VNG1303C

C:Superfamily: conserved hypothetical protein yyaL

Query Match

Best Local Similarity 34.1%; Score 45; DB 2; Length 742;

Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Oy 1 PLHNQEVTAEDGTQRAEKSKPCA 24

Db 701 PLWAGRTGGDGTGPRVVCRRACS 724

|||||:|:|:|:|:|:|

RESULT 21

H96503

protein F9C16.17 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: H96503

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96503

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-873 <STO>

A:Cross-references: UNIPROT:Q9LPO1; UNIPARC:UPI00000AB3E6; GB:AE005173; NID:g8778675; P

C:Genetics:

A:Gene: F9C16.17

A:Map position: 1

Query Match

Best Local Similarity 34.1%; Score 45; DB 2; Length 873;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 5 QEVTAEDGTQRAEKCS 20

Db 83 ESWVADNVVTEKCN 98

:|:|:|:|:|:|

## RESULT 22

S70113

hypothetical protein YDR348c - Yeast (*Saccharomyces cerevisiae*)

A:Alternative names: hypothetical protein D9651.1f

C:Species: *Saccharomyces cerevisiae*

C:Date: 24-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004

C:Accession: S70113

R;Du, Z.

submitted to the EMBL Data Library, March 1996

A:Description: The sequence of *S. cerevisiae* cosmid 9651.

A:Reference number: S70098

A:Accession: S70113

A:Molecule type: DNA

A:Residues: 1-499 &lt;DUZ&gt;

A:Cross-references: UNIPROT:Q05518; UNIPARC:UPI0000053247; EMBL:U51032; NID:gl230659; PI

C:Genetics:

A:Gene: MIPS:YDR348C

A:Cross-references: SGD:S0002756

A:Map position: 4R

C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YDR348c

Query Match

Best Local Similarity 33.7%; Score 44.5; DB 2; Length 499;

Matches 10; Conservative 7; Mismatches 5; Indels 3; Gaps 1;

QY 1 PLHNQEVTAEDGTQRA---EKCSKP 22

DU 164 PPSYEBITSTNGRRRAYPKVKVSRP 188

## RESULT 23

E81715

-ferredoxin [2Fe-2S] TC0329 [similarity] - *Chlamydia muridarum* (strain Nigg)C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C:Accession: E81715

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: E81715

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-91 &lt;TET&gt;

A:Cross-references: UNIPROT:Q9PKY1; UNIPARC:UPI00000578A9; GB:AE002300; GB:AE002160; NID

A:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0329

C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology

C:Keywords: 2Fe, 2S; metalloprotein

F:36,41,44,78/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match

Best Local Similarity 33.3%; Score 44; DB 2; Length 91;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 NQEVTAEDGTQRAEKC 19

DU 12 NQRFLEDGSPIAEAC 27

## RESULT 24

AE3311

cytochrome c1 [imported] - *Brucella melitensis* (strain 16M)C:Species: *Brucella melitensis*

C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

C:Accession: AE3311

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,

.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A:Reference number: AD3252; PMID:11756688

A:Accession: AE3311

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-268 &lt;KUR&gt;

A:Cross-references: UNIPROT:Q8YIG9; UNIPARC:UPI0000057CC6; GB:AE008917; PIDN:AAL51656.1

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10475

A:Map position: 1

C:Superfamily: cytochrome c1 heme protein; cytochrome c1 heme protein homology

C:Keywords: chromoprotein; heme; iron; metalloprotein

F:53.56/Binding site: heme (Cys) (covalent) #status predicted

F:57/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match

Best Local Similarity 33.3%; Score 44; DB 2; Length 268;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGT 13

DU 199 PLSDSQVTYEDGT 211

## RESULT 25

PH1917

FL-160-1 protein - *Trypanosoma cruzi* (fragment)C:Species: *Trypanosoma cruzi*

C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 09-Jul-2004

C:Accession: PH1917

R;Van Voorhis, W.C.; Barrett, L.; Koelling, R.; Farr, A.G.

J. Exp. Med. 178, 681-694, 1993

A:Title: FL-160 proteins of *Trypanosoma cruzi* are expressed from a multigene family and

A:Reference number: JH0823; MUID:93340846; PMID:7688032

A:Accession: PH1917

A:Molecule type: DNA

A:Residues: 1-288 &lt;VAN&gt;

A:Cross-references: UNIPROT:Q03832; UNIPARC:UPI000017B5AE

C:Keywords: glycoprotein

Query Match

Best Local Similarity 33.3%; Score 44; DB 2; Length 288;

Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 HNQEVTAE DGTQRAEKCSP 22

DU 269 HSTVASVDGATGPGNCSR 288

## RESULT 26

T09024

proline-rich protein T27E11.90 - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T09024

R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro,

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533

A:Accession: T09024

A:Molecule type: DNA

A:Residues: 1-577 &lt;BEV&gt;

A:Cross-references: UNIPROT:Q9STP1; UNIPARC:UPI00000A3C75; EMBL:AL078579; GSPDB:GN00062,

A:Experimental source: cultivar Columbia; BAC clone T27E11

C:Genetics:

A:Gene: AtSP.T27E11.90

A:Map position: 4

A:Introns: 26/1, 117/1, 338/1, 411/3, 430/2, 498/2

Query Match

Best Local Similarity 33.3%; Score 44; DB 2; Length 577;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 HNOEVTADGTQRAE 17  
| : ||| : |  
Db 17 HSSNSTASDGTREB 31

## RESULT 27

T40625  
single-stranded DNA-binding protein 68k chain [validated] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40625; T46571  
R:Lyne, M.; Rajandream, M.A.; Barrrell, B.G.; Rieger, M.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z21941  
A:Accession: T40625  
A:Status: preliminary  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-609 <LYN>  
A:Cross-references: UNIPROT:Q92372; UNIPARC:UPI000013379C; EMBL:AL034563; PIDN:CAA22533.  
A:Experimental source: strain 972h-; cosmid c660  
R:Ishai, M.; Sanchez, J.P.; Amin, A.A.; Murakami, Y.; Hurwitz, J.  
J. Biol. Chem. 271, 20868-20878, 1996  
A:Title: Purification, gene cloning and reconstitution of the heterotrimeric single-stranded DNA-binding protein  
A:Reference number: Z23078; MUID:96355433; PMID:8702843  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-609 <ISH>  
A:Cross-references: UNIPARC:UPI000013379C; EMBL:U59385; NID:gl502412; PIDN:AAC49437.1; EMBL:U59385  
C:Genetics:  
A:Gene: SPBC660.l3c; s8b1  
A:Map position: 2  
A:Introns: 14/2  
C:Complex: The single-stranded DNA-binding protein is a heterotrimer, consisting of a 68 kDa subunit and two 28 kDa subunits.  
C:Superfamily: replication protein A1

Query Match 33.3%; Score 44; DB 2; Length 609;  
Best Local Similarity 44.4%; Pred. No. 1.2e+02;  
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 4 NOEVTADGTQRAEKCCK 21  
| : | : | : | : |  
Db 483 NKKVFDQGSWRCCKNK 500

## RESULT 28

T04434  
replication protein A1 homolog T18B16.100 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04434  
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15359  
A:Accession: T04434  
A:Molecule type: DNA  
A:Residues: 1-717 <BEV>  
A:Cross-references: UNIPROT:O49671; UNIPARC:UPI000000A0051; EMBL:AL021687  
A:Experimental source: cultivar Columbia; BAC clone T18B16  
C:Genetics:  
A:Map position: 4  
A:Introns: 96/2; 165/1; 350/3  
A:Note: T18B16.100  
C:Superfamily: replication protein A1

Query Match 33.3%; Score 44; DB 2; Length 717;  
Best Local Similarity 72.7%; Pred. No. 1.5e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 DGTQRAEKCCK 21  
| : | : | : | : |  
Db 483 DGTWRCEKCDK 493

## RESULT 29

C95367  
conserved hypothetical protein SMa1548 [imported] - Sinorhizobium meliloti (strain 1021)  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 16-Aug-2004  
C:Accession: C95367  
R:Barnett, M.J.; Fisher, R.P.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycski, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: C95367  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1051 <KUR>  
A:Cross-references: UNIPROT:Q92YNO; UNIPARC:UPI000000CB20A; GB:AE006469; PIDN:AAK65501.1  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMa1548  
A:Genome: plasmid  
C:Superfamily: Sensor diguanylate cyclase/c-di-GMP phosphodiesterase

Query Match 33.3%; Score 44; DB 2; Length 1051;  
Best Local Similarity 42.1%; Pred. No. 2.1e+02;  
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 5 QEVTADGTQRAEKCCK 23  
| : | : | : | : |  
Db 70 EQITPDGTQRTLLTKRC 88

## RESULT 30

T42636  
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type gamma - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42636  
R:Qinghua, X.; Xiaojun, G.; Cong, S.; Zong, S.M.; Jong, Y.J.; Chan, J.; Wang, L.H.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: Z22226  
A:Accession: T42636  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-1422 <QIN>  
A:Cross-references: UNIPROT:Q98936; UNIPARC:UPI0000132993; EMBL:U38349; NID:gl617477; P  
A:Experimental source: kidney and brain  
C:Superfamily: protein-tyrosine-phosphatase, receptor type gamma; carbonic anhydrase ho  
ne-phosphatase homology  
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane pro

Query Match 33.3%; Score 44; DB 2; Length 1422;  
Best Local Similarity 50.0%; Pred. No. 2.8e+02;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTQRAEK 18  
| : | : | : | : |  
Db 970 PLHGQEHKDEKGSERKPK 987

## RESULT 31

W1WLB5  
E1 protein - human papillomavirus type 5b  
C:Species: human papillomavirus type 5b

```
A>Note: host Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: A40480
R:Yabe, Y.; Sakai, A.; Hisumoto, T.; Kato, H.; Ogura, H.
Virology 183, 793-798, 1991
A:Title: A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic segment amplifi
A:Reference number: A40480; MUID:91306467; PMID:1649510
C:Accession: A40480
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-606 <YAB>
A:Cross-references: UNIPROT:P26542; UNIPARC:UPI0000138316; GB:D90252; NID:G222395; PIDN:
C:Superfamily: papillomavirus E1 protein
C:Keywords: early protein

Query Match      33.0%; Score 43.5; DB 1; Length 606;
Best Local Similarity 55.0%; Pred. No. 1.5e+02;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY      2 LHNQEVTAEDG-TQRAEKCS 20
      ||||| :||| :|||
Db      579 LSDQEEEGDGSQRAFCQS 598

RESULT 32
I52603
MPS1 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
R:Spillsbury, K.; O'Mara, M.A.; Wu, W.M.; Rowe, P.B.; Symonds, G.; Takayama, Y.
Blood 85, 1620-1629, 1995
A:Title: Isolation of a novel macrophage-specific gene by differential cDNA analysis.
A:Reference number: I52603; MUID:95195232; PMID:7888681
A:Accession: I52603
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-661 <RES>
A:Cross-references: UNIPROT:Q61889; UNIPARC:UPI00002910D; GB:L20315; NID:G431419; PIDN:

Query Match      33.0%; Score 43.5; DB 2; Length 661;
Best Local Similarity 43.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

QY      1 PLHNQEVTAEDGTQRAEKCKPC 23
      ||||| :||| :|||
Db      414 PVHLLSQTHEGYSLRE-CKKKC 435

RESULT 33
S60588
drebrin A - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60588
R:Shirao, T.; Kojima, N.; Obata, K.
Neuroreport 3, 109-112, 1992
A:Title: Cloning of drebrin A and induction of neurite-like processes in drebrin-transfe
A:Reference number: S60588; MUID:92305233; PMID:1611026
A:Accession: S60588
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-707 <SHI>
A:Cross-references: UNIPROT:Q07266; UNIPARC:UPI0000129878; EMBL:X59267; NID:G297820; PIDN:

Query Match      33.0%; Score 43.5; DB 2; Length 707;
Best Local Similarity 30.6%; Pred. No. 1.7e+02;
Matches 11; Conservative 5; Mismatches 7; Indels 13; Gaps 1;

QY      2 LHNQEVTAEDGTQRAE-----KCKSPCA 24
      ||||| :||| :|||
Db      637 LTNGETTQEGTQQAASEGYFSQSQBEEFAQSBEPCA 672
```

```
RESULT 34
G71562
ferredoxin [2Fe-2S] IV [similarity] - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71562
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71562
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <ARN>
A:Cross-references: UNIPROT:O84062; UNIPARC:UPI00000D70D8; GB:AE001280; GB:AE001273; NID:
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein
F:23-79/Domain: ferredoxin [2Fe-2S] homology <FER>
F:36,41,44,78/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match      32.6%; Score 43; DB 2; Length 91;
Best Local Similarity 56.2%; Pred. No. 28;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      4 NOEVTAEADGTQRAEKC 19
      ||||| :||| :|||
Db      12 NOEPHLEDGSSIAEVC 27

RESULT 35
I39734
CBS domain-containing protein - Anabaena variabilis
N:Alternate names: hypothetical protein 3
C:Species: Anabaena variabilis
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S68184; I39734
R:Schmitz, O.; Boison, G.; Hilscher, R.; Hundeshagen, B.; Zimmer, W.; Lottspeich, F.; B
Eur. J. Biochem. 233, 266-276, 1995
A:Title: Molecular biological analysis of a bidirectional hydrogenase from cyanobacteri
A:Reference number: I39730; MUID:96061958; PMID:7588754
A:Accession: S68184
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-199 <SCH>
A:Cross-references: UNIPROT:Q44516; UNIPARC:UPI00000B4811; EMBL:X79285; NID:G1032475; P
A:Experimental source: ATCC 29413
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994
C:Superfamily: uncharacterized protein with CBS and CP12 domains; CBS homology
C:Keywords: duplication
F:11-60/Domain: CBS homology <CBS1>
F:78-124/Domain: CBS homology <CBS2>

Query Match      32.6%; Score 43; DB 1; Length 199;
Best Local Similarity 58.8%; Pred. No. 60;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      5 QEVTAEDGTQRAEKCSK 21
      ||||| :||| :|||
Db      167 EEMQAEHQAERAEKVS 183

RESULT 36
AC1902
hypothetical protein alr0765 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1902
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
```

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AC1902  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-204 <KUR>  
A;Cross-references: UNIPROT:Q8YVY1; UNIPARC:UPI00000CDE6C; GB:BA000019; PIDN:BA072722.1;  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr0765  
C;Superfamily: uncharacterized protein with CBS and CPl2 domains; CBS homology

Query Match 32.6%; Score 43; DB 2; Length 204;  
Best Local Similarity 58.8%; Pred. No. 61;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Db 167 EEMQAEIAHQAEKVK 183  
: : || ||||| ||  
: : || ||||| ||

QY 5 QEVTAEDGTQRAEKCSK 21  
: : || ||||| ||  
Db 167 EEMQAEIAHQAEKVK 183

RESULT 37  
T20005  
hypothetical protein C47E12.10 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T20005  
R;Coles, L.  
submitted to the EMBL Data Library, January 1996  
A;Reference number: Z19210  
A;Accession: T20005  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-296 <WIL>  
A;Cross-references: UNIPROT:Q18675; UNIPARC:UPI0000077DD; EMBL:Z68882; PIDN:CAA93102.1;  
A;Experimental source: clone C47E12  
C;Genetics:  
A;Gene: CRSP:C47E12.10  
A;Map position: 4  
A;Introns: 20/3; 88/3; 129/1; 188/3; 221/3

Query Match 32.6%; Score 43; DB 2; Length 296;  
Best Local Similarity 40.0%; Pred. No. 87;  
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 LHNQEVTAEDGTQRAEKCSK 21  
|||: |||: |||: |||:  
Db 193 LHNQEVTAEDGTQRAEKCSK 212

RESULT 38  
JC7266  
3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) 7B - human  
N;Alternate names: cAMP-specific phosphodiesterase 7B  
C;Species: Homo sapiens (man)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: JC7266  
R;Sasaki, T.; Kotera, J.; Yuasa, K.; Omori, K.  
Biochem. Biophys. Res. Commun. 271, 575-583, 2000  
A;Title: Identification of human PDE7B, a cAMP-specific phosphodiesterase.  
A;Reference number: JC7266  
A;Accession: JC7266  
A;Molecule type: mRNA  
A;Residues: 1-450 <SAS>  
A;Cross-references: UNIPROT:Q3NP56; UNIPARC:UPI00000339FF; DBJ:AB038040  
A;Experimental source: caudate nucleus  
C;Genetics:  
A;Gene: pde7B  
A;Map position: 6q23-24  
C;Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',5'-  
C;Keywords: phosphoric diester hydrolase

Query Match 32.6%; Score 43; DB 2; Length 450;  
Best Local Similarity 35.7%; Pred. No. 1.3e+02;  
Matches 10; Conservative 4; Mismatches 8; Indels 6; Gaps 1;

QY 2 LHNQEVTAEDGTQ-----AEKCSKPC 23  
|||: |||: |||: |||:  
Db 298 LHNKDLRLLEDAQRHFLQIALKCADIC 325

RESULT 39  
S28060  
serum response factor-related protein SL2 - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: S28060  
R;Chambers, A.E.; Kotecha, S.; Towers, N.; Mohun, T.J.  
EMBO J. 11, 4981-4991, 1992  
A;Title: Muscle-specific expression of SRF-related genes in the early embryo of Xenopus  
A;Reference number: S28059; MUID:93099873; PMID:1281451  
A;Accession: S28060  
A;Molecule type: mRNA  
A;Residues: 1-516 <CHA>  
A;Cross-references: UNIPROT:Q03414; UNIPARC:UPI000012EEB4; EMBL:Z19123; NID:g65202; PID  
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 32.6%; Score 43; DB 2; Length 516;  
Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PLHNQSVTAEDGTO 14  
||: ||: ||: ||:  
Db 326 PLNTQRTSSQGTQ 339

RESULT 40  
JC6556  
Y chromosome sex-determining-type high mobility group box protein - rainbow trout  
N;Alternate names: Sox23 protein  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004  
C;Accession: JC6504; JC6556  
R;Yamashita, A.; Suzuki, S.; Fujitani, K.; Kojima, M.; Kanda, H.; Ito, M.; Takamatsu, N.  
Gene 209, 193-200, 1998  
A;Title: cDNA cloning of a novel rainbow trout SRY-type HMG box protein, rtSox23, and  
A;Reference number: JC6556; MUID:98192528; PMID:9524264  
A;Accession: PC6504  
A;Molecule type: protein  
A;Residues: 174-585 <YAZ>  
A;Cross-references: UNIPROT:O57396; UNIPARC:UPI000017BFSF  
A;Experimental source: ovary  
A;Accession: JC6556  
A;Molecule type: mRNA  
A;Residues: 1-585 <YAM>  
A;Cross-references: UNIPARC:UPI00000FB858; DBJ:AB007906; NID:g2789467; PIDN:BA024402.1;  
C;Comment: This protein forms a homodimer through the leucine zipper motif.  
C;Genetics:  
A;Gene: rtSox23  
C;Keywords: homodimer; leucine zipper; ovary  
F;429-504/Domain: HMG box homology <HMG1>

Query Match 32.6%; Score 43; DB 2; Length 585;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 QEVTAEDGTQRAEKCSK 22  
|||: |||: |||: |||:  
Db 28 QESGGGGAKAEPCCSP 45

Search completed: December 3, 2005, 07:50:43  
Job time : 30.0909 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 3, 2005, 07:35:08 ; Search time 159.273 Seconds  
(without alignment)  
106.312 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLNNQEVTAEDGTQAEKCKPKCA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	419	2	Q9UK79 HUMAN
2	128	97.0	1255	1	ERBB2 HUMAN
3	121	91.7	1259	1	ERBB2_CANFA
4	114	86.4	711	2	Q80R89_MOUSE
5	114	86.4	1257	1	ERBB2_RAT
6	114	86.4	1259	2	Q8K3F9_RAT
7	114	86.4	1305	2	Q6ZPE0_MOUSE
8	113	85.6	1254	1	ERBB2_MBSAU
9	55	41.7	1195	2	Q4RV10_TETNG
10	54	40.9	92	1	SGPI_SCHGR
11	52	39.4	273	1	OLRI_HUMAN
12	51	38.6	498	2	Q7S109_NEUCR
13	50.5	38.3	670	2	Q4QG99_LEIMA
14	50	37.9	676	2	Q59P42_HUMAN
15	50	37.9	686	1	2NF7_HUMAN
16	50	37.9	686	2	Q5R8X0_PONPY
17	50	37.9	944	2	Q6CNE1_KLUJA
18	49	37.1	74	2	Q9N0X1_RABIT
19	49	37.1	116	2	Q9BG67_RABIT
20	49	37.1	147	1	VXIS_BPSF5
21	49	37.1	405	2	Q6RGS5_HUMAN
22	49	37.1	645	2	Q6RGS6_HUMAN
23	49	37.1	1081	2	Q59F18_HUMAN
24	49	37.1	1091	2	Q504U8_HUMAN
25	49	37.1	1209	2	Q8MIL8_PIG
26	49	37.1	1210	1	EGFR_HUMAN
27	49	37.1	1913	2	Q52PE0_MAGGR
28	48.5	36.7	324	2	Q90XJ6_LATWE
29	48.5	36.7	514	2	Q6JQ08_LATWE
30	48.5	36.7	912	2	Q6T1H0_LATWE
31	48	36.4	103	2	Q5M840_RAT

32	48	36.4	150	2	Q6IGT7_DROME
33	48	36.4	214	2	Q8XDI3_ECO57
34	48	36.4	219	2	Q83LE8_SHIFL
35	48	36.4	270	2	Q87503_ECOLI
36	48	36.4	468	2	Q4QBD3_LEIMA
37	48	36.4	478	2	Q9ESE0_RAT
38	48	36.4	632	2	Q6NKT4_ARATH
39	48	36.4	643	2	Q9ERV6_MOUSE
40	48	36.4	644	2	Q23J21_ARATH
41	48	36.4	655	2	Q9WVF5_MOUSE
42	48	36.4	882	2	Q6XOM1_9STRA
43	48	36.4	1209	2	Q9QX70_RAT
44	48	36.4	1210	1	EGFR_MOUSE
45	48	36.4	1210	2	Q9EP98_MOUSE
46	48	36.4	1210	2	Q5SVE8_MOUSE
47	48	36.4	1224	2	Q93330_CAEEL
48	48	36.4	1435	1	TRS85_HUMAN
49	48	36.4	1445	2	Q4QG46_LEIMA
50	47.5	36.0	449	2	Q5XGJ3_XENTR
51	47.5	36.0	482	2	Q7RE84_PLAYO
52	47	35.6	64	2	Q615E1_CAEBR
53	47	35.6	280	2	Q5BIC0_EMENT
54	47	35.6	500	2	Q4K516_PSEF5
55	47	35.6	572	2	Q9AU64_ELAGV
56	47	35.6	573	2	Q8BZU6_MOUSE
57	47	35.6	655	1	TNR21_MOUSE
58	47	35.6	655	2	Q543Y9_MOUSE
59	47	35.6	684	2	Q7Q2B5_ANOGA
60	47	35.6	722	2	Q9DC42_MOUSE
61	47	35.6	772	2	Q6MYK9_ASPFU
62	47	35.6	774	2	Q80379_DAUCA
63	47	35.6	777	2	Q7QVA9_GIALA
64	47	35.6	811	2	Q4QB17_LEIMA
65	47	35.6	987	2	Q8MR93_DROME
66	47	35.6	987	2	Q9V7X1_DROME
67	46.5	35.2	77	2	Q9AR26_AVIMR
68	46.5	35.2	284	2	Q4H3J8_CIOIN
69	46.5	35.2	450	2	Q7X2U6_9BACT
70	46	34.8	66	2	Q6V7P8_9CAUD
71	46	34.8	184	2	Q98NF4_RHILU
72	46	34.8	195	2	Q96VJ6_KLUZA
73	46	34.8	224	2	Q4IA61_GIBZE
74	46	34.8	314	2	Q5AK03_CANAL
75	46	34.8	580	2	Q6F756_ACIAD
76	46	34.8	651	2	Q98SM6_CHICK
77	46	34.8	662	2	Q5X128_BRARE
78	46	34.8	678	2	Q5REJ3_PONPY
79	46	34.8	937	2	Q6Z3H8_ORYSA
80	46	34.8	963	2	Q7XL40_ORYSA
81	46	34.8	974	2	Q8LNB2_ORYSA
82	46	34.8	1195	2	Q7XQZ1_ORYSA
83	46	34.8	1390	1	INSR_AEDAE
84	46	34.8	1449	2	Q8S5K3_ORYSA
85	46	34.8	1461	2	Q7XGE3_ORYSA
86	46	34.8	1466	2	Q7XGE5_ORYSA
87	46	34.8	1661	2	Q7XBQ6_ORYSA
88	46	34.8	2647	2	Q90580_PLAFA
89	45.5	34.5	436	2	Q4Z8N7_PIG
90	45.5	34.5	474	2	Q8G4D3_BIFLO
91	45.5	34.5	704	2	Q9DA98_MOUSE
92	45.5	34.5	709	2	Q99P25_MOUSE
93	45	34.1	174	2	Q8SA72_ORYSA
94	45	34.1	179	2	Q6B8H4_9ACAR
95	45	34.1	196	2	Q7RUC9_NEUCR
96	45	34.1	205	2	Q26554_SCHMA
97	45	34.1	231	2	Q8K5B9_MOUSE
98	45	34.1	264	2	Q7XBH7_ORYSA
99	45	34.1	289	2	Q6FRM9_CANGA
100	45	34.1	303	2	Q8TWD2_METKA

ALIGNMENTS

```
RESULT 1
Q9UK79 HUMAN
ID Q9UK79_HUMAN PRELIMINARY; PRT; 419 AA.
AC Q9UK79
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Herstatin.
GN Name=HER-2;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99415951; PubMed=10485918; DOI=10.1073/pnas.96.19.10869;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177761; AAD56009.2; -; mRNA.
DR HSSP; P04626; 1N8Z.
DR SNR; Q9UK79; 23-345.
DR Ensemble; ENSG00000141736; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IDA.
DR GO; GO:0004715; F:non-membrane spanning protein tyrosine kina. ; IDA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. ; IDA.
DR InterPro; IPR000494; EGRF_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin_repeat.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 1.
DR SMART; SM00261; FU; 1.
SQ SEQUENCE 419 AA; 45472 MW; FEC1BE347E2D030C CRC64;

Query Match 97.0%; Score 128; DB 2; Length 419;
Best Local Similarity 95.8%; Pred. No. 4;le-11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLHNQVETADGTQRAEKCSPCA 24
| | | | | | | | | | | | | | | | | | | |
Db 316 PLHNQVETADGTQRAEKCSPCA 339

RESULT 2
ERBB2 HUMAN
ID ERBB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626; Q14256; Q6LDV1; Q9UMK4;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbB-2 precursor (BC 2.7.1.112)
DE (p18erBB2) (C-erbB-2) (NEU proto-oncogene) (Tyrosine kinase-type cell
DE surface receptor HER2) (MLN 19).
GN Name=ERBB2; Synonyms=HER2, NEU, NGL;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,
RA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
```

```
RL Nature 319:230-234 (1986).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Libermann T.A., Schlessinger J., Francke U.,
RA Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139 (1985).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS CYS-452; VAL-655 AND
RP ALA-1170.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-191.
RX MEDLINE=87286898; PubMed=3039351;
RA Tal M., King C.R., Kraus M.H., Ullrich A., Schlessinger J., Givol D.;
RT "Human HER2 (neu) promoter: evidence for multiple mechanisms for
RT transcriptional initiation.";
RL Mol. Cell. Biol. 7:2597-2601 (1987).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 737-1031.
RX MEDLINE=86016729; PubMed=2995967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-
RT erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501 (1985).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 832-909.
RX TISSUE=Mammary carcinoma;
RX MEDLINE=85272597; PubMed=2992089;
RA King C.R., Kraus M.H., Aaronson S.A.;
RT "Amplification of a novel v-erbB-related gene in a human mammary
RT carcinoma.";
RL Science 229:974-976 (1985).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 1081-1245, AND VARIANT ALA-1170.
RX MEDLINE=94000386; PubMed=8104414;
RA Sarkar F.H., Ball D.E., Li Y.W., Criseman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
RT gene.";
RL DNA Cell Biol. 12:611-615 (1993).
RN [8]
RP IDENTIFICATION IN A COMPLEX WITH PIK3C2A AND EGFR, IDENTIFICATION IN A
RP COMPLEX WITH PIK3C2B AND EGFR, AND INTERACTION WITH PIK3C2B.
RX PubMed=10805725;
RA Arcaro A., Zvelebil M.J., Wallasch C., Ullrich A., Waterfield M.D.,
RA Domin J.;
RT "Class II phosphoinositide 3-kinases are downstream targets of
RT activated polypeptide growth factor receptors.";
RL Mol. Cell. Biol. 20:3817-3830 (2000).
RN [9]
RP INTERACTION WITH PLXNB1.
RX PubMed=15210733; DOI=10.1083/jcb.200312094;
RA Swiercz J.M., Kumer R., Offermanns S.;
RT "Plexin-B1/RhoGEF-mediated RhoA activation involves the receptor
RT tyrosine kinase ErbB-2.";
RL J. Cell Biol. 165:869-880 (2004).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 654-662 IN COMPLEX WITH HLA
RP AND BETA-2 MICROGLOBULIN.
RX MEDLINE=20062861; PubMed=10593938; DOI=10.1074/jbc.274.51.36422;
RA Kuhs J.J., Batalia M.A., Yan S., Collins E.J.;
RT "Poor binding of a HER-2/neu epitope (GP2) to HLA-A2.1 is due to a
RT lack of interactions with the center of the peptide.";
RL J. Biol. Chem. 274:36422-36427 (1999).
```



CC PIK3C2B when phosphorylated on Tyr-1200 (By similarity).  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -|- PTM: Ligand-binding increases phosphorylation on tyrosine residues  
CC (By similarity).  
CC -|- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor  
CC subfamily.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; AB008451; BAA23127.1; -; mRNA.  
CC DR HSSP; P04626; 1N82.  
CC DR SMR; O18735; 23-628.  
CC DR Ensembl; ENSCARG00000016351; Canis familiaris.  
CC DR InterPro; IPR000494; EGFR\_L.  
CC DR InterPro; IPR006211; Furin-like.  
CC DR InterPro; IPR006212; Furin repeat.  
CC DR InterPro; IPR000719; Prot Kinase.  
CC DR InterPro; IPR001245; Tyr\_Pkinase.  
CC DR InterPro; IPR008266; Tyr\_Pkinase\_AS.  
CC DR InterPro; IPR004019; YLP\_motif.  
CC DR Pfam; PF00757; Furin-like; 1.  
CC DR Pfam; PF02757; Recep\_L domain; 2.  
CC DR PRINTS; PR00109; TYRKINASE.  
CC DR ProDom; PD000001; Prot\_kinase; 1.  
CC DR SMART; SM00261; FU; 3.  
CC DR SMART; SM00219; TyrKc; 1.  
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
CC DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC KW ATP-binding; Glycoprotein; Kinase; Multigene family;  
KW Nucleotide-binding; Phosphorylation; Receptor; Signal; Transferase;  
KW Transmembrane; Tyrosine-protein kinase.  
KW SIGNAL 1 22 Potential.  
FT CHAIN 23 1259 Receptor tyrosine-protein kinase erbB-2.  
FT FT TOPO\_DOM 23 653 Extracellular (Potential).  
FT FT TRANSMEM 654 674 Potential.  
FT FT TOPO\_DOM 675 1259 Cytoplasmic (Potential).  
FT FT DOMAIN 719 986 Protein kinase.  
FT FT NP\_BIND 725 1201 ATP (By similarity).  
FT FT REGION 1199 1201 Interaction with PIK3C2B (By similarity).  
FT FT COMPIAS 192 268 Cys-rich.  
FT FT COMPIAS 1101 1219 Pro-rich.  
FT FT ACT\_SITE 844 844 By similarity.  
FT FT BINDING 752 752 ATP (By similarity).  
FT FT MOD\_RES 1138 1138 Phosphotyrosine (by autocatalysis) (By  
FT similarity).  
FT MOD\_RES 1200 1200 Phosphotyrosine (Potential).  
FT FT MOD\_RES 1252 1252 Phosphotyrosine (by autocatalysis) (By  
FT similarity).  
FT CARBOHYD 68 68 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 259 259 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 421 421 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 529 529 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 570 570 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 628 628 N-linked (GlcNAc... ) (Potential).  
FT DISULFID 195 204 By similarity.  
FT DISULFID 199 212 By similarity.  
FT DISULFID 220 227 By similarity.  
FT DISULFID 224 235 By similarity.  
FT DISULFID 236 244 By similarity.  
FT DISULFID 240 252 By similarity.  
FT DISULFID 255 264 By similarity.  
FT DISULFID 268 295 By similarity.  
FT DISULFID 299 311 By similarity.  
FT DISULFID 315 331 By similarity.  
FT DISULFID 334 338 By similarity.  
FT DISULFID 511 519 By similarity.  
FT DISULFID 514 527 By similarity.

FT DISULFID 530 539 By similarity.  
FT DISULFID 543 559 By similarity.  
FT DISULFID 562 575 By similarity.  
FT DISULFID 586 595 By similarity.  
FT DISULFID 599 622 By similarity.  
FT DISULFID 625 633 By similarity.  
FT DISULFID 629 641 By similarity.  
SQ SEQUENCE 1259 AA; 137991 MW; E37364D49C4ACD46 CRC64;  
Query Match 91.7%; Score 121; DB 1; Length 1259;  
Best Local Similarity 91.7%; Pred. No. 1.6e-09;  
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLNQEVTAEDGTQRAEKCKSPCA 24  
DB 316 PLNQEVTAEDGTQRAEKCKSPCA 339  
-----  
RESULT 4  
Q80Y89 MOUSE  
ID Q80Y89\_MOUSE PRELIMINARY; PRT; 711 AA.  
AC Q80Y89;  
CT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE V-erb-b2 erythroblastic leukemia viral oncogene homolog 2,  
DE neuro/glioblastoma derived oncogene homolog, isoform 2.  
GN Name=ErbB2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.J., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M.J., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield A.S., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RG NIH MGC Project;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RG NIH MGC Project;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC046811; AAH46811.1; -; mRNA.  
DR EMBL; BC053078; AAH53078.1; -; mRNA.  
DR HSSP; P06494; 1N8Y.  
DR SMR; Q80Y89; 23-630.  
DR Ensembl; ENSMUSG000000062312; Mus musculus.

```

DR MGI; MGI:954110; Erbb2.
DR GO; GO:0016324; C:apical plasma membrane; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007507; P:heart development; IMP.
DR GO; GO:0042552; P:myelination; IMP.
DR GO; GO:0007422; P:peripheral nervous system development; IMP.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin-repeat.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 711 AA; 78707 MW; 682B188EB0E71318 CRC64;

Query Match      86.48; Score 114; DB 2; Length 711;
Best Local Similarity 87.58; Pred.No. 1.1e-08;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PLNQEVTAEDGTQRAEKCCKPCA 24
   | : ||||| ||||| ||||| |||||
Db 317 PPNQEVTAEDGTQRCCKCKPCA 340

RESULT 5
ERBB2 RAT
ID ERBB2 RAT STANDARD; PRT; 1257 AA.
AC P06494; Q6P732;
DT 01-JAN-1998 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene) (Epidermal growth factor
DE receptor-related protein).
DE Name=Erbb2; Synonyms=Neu;
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Neuroblastoma;
RX MEDLINE=86118662; PubMed=3945311;
RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
RT "The neu oncogene encodes an epidermal growth factor receptor-related
RT protein.";
RL Nature 319:226-230(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE OF 634-699.
RX MEDLINE=92035293; PubMed=1682063;
RA Masui T., Mann A.M., Macatee T.L., Garland E.M., Okamura T.,
RA Smith R.A., Cohen S.M.;
RT "Direct DNA sequencing of the rat neu oncogene transmembrane domain
RT reveals no mutation in urinary bladder carcinomas induced by N-butyl-
RT N-(4-hydroxybutyl)nitrosamine, N-[4-(5-nitro-2-furyl)-2-
RT thiazolyl]formamide or N-methyl-N-nitrosourea.";
RL Carcinogenesis 12:1975-1978(1991).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 852-905.
RC TISSUE=Sciatic nerve;
RX MEDLINE=91222560; PubMed=2025425; DOI=10.1016/0896-6273(91)90167-X;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system.";
RL Neuron 6:691-704(1991).
RN [5]
RP STRUCTURE BY NMR OF 650-668.
RX MEDLINE=92155181; PubMed=1346763;
RA Gullick W.J., Bottomley A.C., Lofts F.J., Doak D.G., Mulvey D.,
RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
RT "Three dimensional structure of the transmembrane region of the proto-
RT oncogenic and oncogenic forms of the neu protein.";
RL EMBO J. 11:43-48(1992).
CC -!- FUNCTION: Essential component of a neurotulin-receptor complex,
CC although neurotulin do not interact with it alone. GP30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-
CC alpha and amphiregulin.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein
CC tyrosine phosphate.
CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
CC constitutively activated oncogenic variant forms a homodimer.
CC Interacts with PRKCAP and PLXNB1. Part of a complex with EGFR and
CC either PIK3C2A or PIK3C2B. May interact with PIK3C2B when
CC phosphorylated on Tyr-1198 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine residues
CC (By similarity).
CC -!- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor
CC subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X03362; CAA27059.1; ALT_INIT; mRNA.
CC EMBL; BC061863; AAH61863.1; ALT_INIT; mRNA.
CC PIR; A24562; TVRTNU.
CC PDB; 1N8Y; X-ray; C=23-631.
CC Ensembl; ENSRNOG00000006450; Rattus norvegicus.
CC RCD; 2561; Erbb2.
CC GO; GO:0045595; P:regulation of cell differentiation; TAS.
CC GO; GO:0042127; P:regulation of cell proliferation; TAS.
CC InterPro; IPR000494; EGFR_L.
CC InterPro; IPR006211; Furin-like.
CC InterPro; IPR006212; Furin-repeat.
CC InterPro; IPR00719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC InterPro; IPR004019; YLP_motif.
CC Pfam; PF00757; Furin-like; 1.
CC Pfam; PF01030; Recep_L_domain; 2.
CC Pfam; PF02757; YLP; 2.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00261; FU; 4.
CC SMART; SM00219; TyrcK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC 3D-structure; ATP-binding; Disease mutation; Glycoprotein; Kinase;
KW Multigene family; Nucleotide-binding; Phosphorylation; Proto-oncogene;
KW Receptor; Signal; Transferase; Transmembrane; Tyrosine-protein kinase.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 1257 Receptor tyrosine-protein kinase erbB-2.
FT TOPO_DOM 22 654 Extracellular (Potential).
FT TRANSMEM 655 677 Potential.
FT TOPO_DOM 678 1257 Cytoplasmic (Potential).
FT DOMAIN 722 989 Protein kinase.
FT NP_BIND 728 1197 ATP (By similarity).
FT REGION 1197 1199 Interaction with PIK3C2B (By similarity).
FT COMEBIAS 159 369 Cys-rich.
FT COMEBIAS 473 646 Cys-rich.
FT ACT_SITE 847 847 By similarity.
FT BINDING 755 755 ATP (By similarity).
FT MOD_RES 1141 1141 Phosphotyrosine (by autocatalysis) (By
FT MOD_RES 1198 1198 Phosphotyrosine (Potential).
FT MOD_RES 1250 1250 Phosphotyrosine (by autocatalysis) (By

```

```

FT CARBOHYD 68 68 similarity).
FT CARBOHYD 188 188 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 260 260 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 532 532 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 573 573 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 631 631 N-linked (GlcNAc... ) (Potential).
FT DISULFID 196 205 By similarity.
FT DISULFID 200 213 By similarity.
FT DISULFID 221 228 By similarity.
FT DISULFID 225 236 By similarity.
FT DISULFID 237 245 By similarity.
FT DISULFID 241 253 By similarity.
FT DISULFID 256 265 By similarity.
FT DISULFID 269 296 By similarity.
FT DISULFID 300 312 By similarity.
FT DISULFID 316 332 By similarity.
FT DISULFID 335 339 By similarity.
FT DISULFID 513 522 By similarity.
FT DISULFID 517 530 By similarity.
FT DISULFID 533 542 By similarity.
FT DISULFID 546 562 By similarity.
FT DISULFID 565 578 By similarity.
FT DISULFID 569 586 By similarity.
FT DISULFID 589 598 By similarity.
FT DISULFID 602 625 By similarity.
FT DISULFID 628 636 By similarity.
FT DISULFID 632 644 By similarity.
FT VARIANT 661 661 V -> E (in oncogenic NEU).
FT CONFLICT 145 145 S -> G (in Ref. 2).
FT CONFLICT 505 509 LCVSS -> CGUE (in Ref. 2).
SQ SEQUENCE 1257 AA; 138932 MW; 6129264593011402 CRC64;

Query Match 86.4%; Score 114; DB 1; Length 1257;
Best Local Similarity 87.5%; Pred. No. 2e-08; 2; Indels 0; Gaps 0;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
| : ||||| ||||| |||||
Db 317 PPNNQEVTAEDGTQRCCKSKPCA 340

RESULT 6
Q8K3F9 RAT PRELIMINARY; PRT; 1259 AA.
AC Q8K3F9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
MT 01 OCT 2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Neu protooncoprotein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BDIX;
RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY116182; AA050093.1; -; mRNA.
DR HSSP; P06494; 1N8Y.
DR SMR; Q8K3F9; 26-633.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin...; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008366; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.

```

```

DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 1259 AA; 139101 MW; B724BD5CC33AE953 CRC64;

Query Match 86.4%; Score 114; DB 2; Length 1259;
Best Local Similarity 87.5%; Pred. No. 2e-08; 2; Indels 0; Gaps 0;
Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLHNQEVTAEDGTQRAEKCKSPCA 24
| : ||||| ||||| |||||
Db 320 PPNNQEVTAEDGTQRCCKSKPCA 343

RESULT 7
Q6ZPE0 MOUSE PRELIMINARY; PRT; 1305 AA.
AC Q6ZPE0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE KIAA3023 protein (Fragment).
GN Name=Exrb2; Synonyms=mkIAA3023;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryonic tail;
RX MEDLINE=22977043; PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT III, the complete nucleotide sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
DR EMBL; AK129487; BAC98297.1; -; mRNA.
DR HSSP; P11362; 1AGW.
DR SMR; Q6ZPE0; 72-679.
DR Ensemble; ENSMUSG00000062312; Mus musculus.
DR MGI; MGI:95410; Erbb2.
DR GO; GO:0016324; C:apical plasma membrane; IDA.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007507; F:heart development; IMP.
DR GO; GO:0042552; F:myelination; IMP.
DR GO; GO:0007422; F:peripheral nervous system development; IMP.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008366; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.

```

DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SMO0261; FU; 4.  
 DR SMART; SMO0220; S\_TK; 1.  
 DR SMART; SMO0219; TyrKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; Nucleotide-binding; Transferase.  
 FT NON TER 1  
 SQ SEQUENCE 1305 AA; 143508 MW; A51D897408521860 CRC64;

Query Match 86.4%; Score 114; DB 2; Length 1305;  
 Best Local Similarity 87.5%; Pred. No. 2.1e-08;  
 Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 PLHNOEVTAEADGTORAEKCSKPCA 24  
 DB 366 PPNQOEVTAEADGTORCEKCSKPCA 389

RESULT 8

ERBB2 MESAU STANDARD; PRT; 1254 AA.  
 ID ERBB2 MESAU STANDARD; PRT; 1254 AA.  
 AC Q60553;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Receptor tyrosine-protein kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene).  
 GN Name=ERBB2; Synonym=NEU;  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Cricetidae; Cricetinae; Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Nerve;  
 RX MEDLINE=94193007; PubMed=7908275; DOI=10.1016/0378-1119(94)90553-3;  
 RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M., Yamazaki Y.,  
 RA Ishikawa T.;  
 RA "Cloning and activation of the Syrian hamster neu proto-oncogene.";  
 RL Gene 140:251-255(1994).  
 CC -!- FUNCTION: Essential component of a neurotrophin-receptor complex,  
 CC although neurotrophins do not interact with it alone. GP30 is a  
 CC potential ligand for this receptor. Not activated by EGF, TGF-  
 CC alpha and amphiregulin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
 CC tyrosine phosphate.  
 CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors  
 CC (Potential). Part of a complex with EGFR and either PIK3C2A or  
 CC PIK3C2B. Interacts with PRKCA and PLXNB1. May interact with  
 CC PIK3C2B when phosphorylated on Tyr-1196 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine  
 CC residues.  
 CC -!- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor  
 CC subfamily.

-----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; D16295; BAA03801.1; -; mRNA.  
 CC PIR; I48161; I48161.  
 CC HSP; P06494; IN8Y.  
 CC SMR; Q60553; 23-629.  
 DR InterPro; IPR000494; EGFR\_L.  
 DR InterPro; IPR006211; Furin-like.  
 DR InterPro; IPR006212; Furin repeat.  
 DR InterPro; IPR000719; Prot\_Kinase.

DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR InterPro; IPR004019; YLP\_motif.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR Pfam; PF02757; YLP; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SMO0261; FU; 4.  
 DR SMART; SMO0219; TyrKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Disease mutation; Glycoprotein; Kinase; Multigene family;  
 KW Nucleotide-binding; Phosphorylation; Proto-oncogene; Receptor; Signal;  
 KW Transferase; Transmembrane; Tyrosine-protein kinase.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1254  
 FT TOPO\_DOM 22 652  
 FT TRANSMEM 653 675  
 FT TOPO\_DOM 676 1254  
 FT DOMAIN 720 987  
 FT NP\_BIND 726 734  
 FT REGION 1195 1197  
 FT COMPBIAS 158 368  
 FT COMPBIAS 472 644  
 FT ACT\_SITE 845 845  
 FT BINDING 753 753  
 FT MOD\_RES 1139 1139  
 FT MOD\_RES 1196 1196  
 FT MOD\_RES 1247 1247  
 FT CARBOHYD 68 68  
 FT CARBOHYD 125 125  
 FT CARBOHYD 187 187  
 FT CARBOHYD 259 259  
 FT CARBOHYD 530 530  
 FT CARBOHYD 571 571  
 FT CARBOHYD 629 629  
 FT DISULFID 195 204  
 FT DISULFID 199 212  
 FT DISULFID 236 244  
 FT DISULFID 240 252  
 FT DISULFID 255 264  
 FT DISULFID 268 295  
 FT DISULFID 299 311  
 FT DISULFID 315 331  
 FT DISULFID 334 338  
 FT DISULFID 511 520  
 FT DISULFID 515 528  
 FT DISULFID 531 540  
 FT DISULFID 544 560  
 FT DISULFID 563 576  
 FT DISULFID 567 584  
 FT DISULFID 587 596  
 FT DISULFID 600 623  
 FT DISULFID 626 634  
 FT DISULFID 630 642  
 FT VARIANT 658 658 V -> E (in oncogenic NEU).  
 FT VARIANT 659 659 V -> E (in oncogenic NEU).  
 SQ SEQUENCE 1254 AA; 138253 MW; 974C3791C21P2BE1 CRC64;

Query Match 85.6%; Score 113; DB 1; Length 1254;  
 Best Local Similarity 87.5%; Pred. No. 2.8e-08;  
 Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 PLHNOEVTAEADGTORAEKCSKPCA 24  
 DB 316 PLHNOEVTAEADGTORCEKCSKPCA 339

RESULT 9

```

Q4RV10 TETNG
ID Q4RV10 TETNG PRELIMINARY; PRT; 1195 AA.
AC Q4RV10;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Chromosome 12 SCAF14993, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00028560001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Fritsch J., Gardier S., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEE01014993; CAG07772.1; -; Genomic_DNA.
FT NON_TER 1 1
SQ SEQUENCE 1195 AA; 13325 MW; 52A41A36B3F6605E CRC64;

Query Match 41.7%; Score 55; DB 2; Length 1195;
Best Local Similarity 46.4%; Pred. No. 28;
Matches 13; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 1 PLHNQSVTAEDGTQRAEK-----CSKPCA 24
DB 961 PVNNNNVIQBDTYEWALKSWSPCSKPCA 988

RESULT 10
SGPI_SCHGR
ID SGPI_SCHGR STANDARD; PRT; 92 AA.
AC O46162;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serine protease inhibitor I/II precursor [Contains: Protease inhibitor
DE SGPI-1 (Schistocerca gregaria trypsin inhibitor) (SGPI); Protease
DE inhibitor SGPI-2 (Schistocerca gregaria chymotrypsin inhibitor)
DE (SGCI)].
OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OX NCBI_TaxID=7010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE Ovary;
RX MEDLINE=98314507; PubMed=9652399;
RA Vanden Broeck J., Chiou S.-J., Schoofs L., Hamdaoui A.,

RA
RT "Cloning of two cDNAs encoding three small serine protease inhibiting
RT peptides from the desert locust Schistocerca gregaria and analysis of
RT tissue-dependent and stage-dependent expression.";
RL Eur. J. Biochem. 254:90-95(1998).
RN [2]
RP PROTEIN SEQUENCE OF 20-54 AND 57-91, FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=98133320; PubMed=9475173; DOI=10.1016/S0014-5793(97)01585-8;
RA Hamdaoui A., Wataleb S., Devreese B., Chiou S.-J., Vanden Broeck J.,
RA Van Beumen J., De Loof A., Schoofs L.;
RT "Purification and characterization of a group of five novel peptide
RT serine protease inhibitors from ovaries of the desert locust,
RT Schistocerca gregaria.";
RL FEBS Lett. 422:74-78(1998).
RN [3]
RP STRUCTURE BY NMR OF 20-54 AND 57-91.
RX PubMed=11997226; DOI=10.1016/S1096-4959(01)00530-9;
RA Simonet G., Claeys I., Vanden Broeck J.;
RT "Structural and functional properties of a novel serine protease
RT inhibiting peptide family in arthropods.";
RL Comp. Biochem. Physiol. 132B:247-255(2002).
CC -! FUNCTION: In vitro, SGPI-1/SGCI is active against alpha-
CC chymotrypsin and trypsin while SGPI-2/SGTI is active against
CC alpha-chymotrypsin and pancreatic elastase.
CC -! SUBCELLULAR LOCATION: Secreted.
CC -! TISSUE SPECIFICITY: Expressed in hemolymph, ovaries, testes and
CC fat body of adults but are absent in the gut. Also present in
CC larval hemolymph and fat body.
CC -! SIMILARITY: Belongs to the protease inhibitor I19 family.
CC -! SIMILARITY: Contains 2 pacifastin repeats.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; Y09605; CAA70818.1; -; mRNA.
DR PDB; 1KGM; NMR; A=57-91.
DR PDB; 1KIO; NMR; A=57-91.
DR InterPro; IPR008037; Prot_inh_PMP.
DR Pfam; PF05375; Pacifastin_1; 2.
KW 3D-structure; Cleavage on pair of basic residues;
KW Direct protein sequencing; Protease inhibitor; Repeat;
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 19
FT PEPTIDE 20 54 Protease inhibitor SGPI-1.
FT PEPTIDE 57 91 Protease inhibitor SGPI-2.
FT REPEAT 22 61 Pacifastin 1.
FT REPEAT 62 92 Pacifastin 2.
FT SITE 48 49 Reactive bond.
FT SITE 86 87 Reactive bond.
FT DISULFID 23 38
FT DISULFID 33 51
FT DISULFID 36 46
FT DISULFID 60 75
FT DISULFID 70 89
FT DISULFID 73 84
SQ SEQUENCE 92 AA; 9842 MW; 9E5A228C767C657C CRC64;

Query Match 40.9%; Score 54; DB 1; Length 92;
Best Local Similarity 50.0%; Pred. No. 2.6;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PLHNQSVTAEDGTQRAEKCS 20
DB 52 PPKREVTCPEGTTFDKCN 71

RESULT 11
OLR1_HUMAN
ID OLR1_HUMAN STANDARD; PRT; 273 AA.

```



AC P78380; Q72484;  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Oxidized low-density lipoprotein receptor 1 (Ox-LDL receptor 1)  
DE (Lectin-type oxidized LDL receptor 1) (Lectin-like oxidized LDL  
DE receptor 1) (Lectin-like oxLDL receptor 1) (LOX-1) [Contains:  
DE Oxidized low-density lipoprotein receptor 1, soluble form].  
CN Name=OLR1; Synonyms=LOX1;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.  
RC TISSUE=Lung;  
RX MEDLINE=97205278; PubMed=9052782;  
RA Sawamura T., Kume N., Aoyama T., Moriaki H., Hoshikawa H., Aiba Y.,  
RA Tanaka T., Miwa S., Katsura Y., Kita T., Masaki T.;  
RT "An endothelial receptor for oxidized low-density lipoprotein.";  
RL Nature 386:73-77(1997).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RX MEDLINE=98438416; PubMed=9763655;  
RA Li X., Bouzyk M.M., Wang X.;  
RT "Assignment of the human oxidized low-density lipoprotein receptor  
RT gene (OLR1) to chromosome 12p13.1--p12.3, and identification of a  
RT polymorphic CA-repeat marker in the OLR1 gene.";  
RL Cytogenet. Cell Genet. 82:34-36(1998).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA], DOMAIN, DISULFIDE BONDS, N-GLYCOSYLATION,  
RP AND MUTAGENESIS OF CYS-144; CYS-155; CYS-172; ASN-183;  
RP 209-ARG-ASN-210; HIS-226; ARG-229; ARG-231; 235-SER-GLN-236; SER-240;  
RP CYS-243; CYS-256; CYS-264 AND 267-LYS--GLN-273.  
RX MEDLINE=21157417; PubMed=11256994;  
RA Shi X., Nimi S., Ohtani T., Machida S.;  
RT "Characterization of residues and sequences of the carbohydrate  
RT recognition domain required for cell surface localization and ligand  
RT binding of human lectin-like oxidized LDL receptor.";  
RL J. Cell Sci. 114:1273-1282(2001).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], TISSUE SPECIFICITY, AND INDUCTION.  
RX MEDLINE=99047525; PubMed=9828121; DOI=10.1006/geno.1998.5561;  
RA Yamanaka S., Zhang X.-Y., Miura K., Kim S., Iwao H.;  
RT "The human gene encoding the lectin-type oxidized LDL receptor (OLR1)  
RT is a novel member of the natural killer gene complex with a unique  
RT expression profile.";  
RL Genomics 54:191-199(1998).  
RN [5]  
RP NUCLEOTIDE SEQUENCE [MRNA], INVOLVEMENT IN MYOCARDIAL INFARCTION, AND  
RP VARIANT ASN-167.  
RX MEDLINE=22533240; PubMed=12646194;  
RA Tatsuguchi M., Furutani M., Hinagata J.-I., Tanaka T., Furutani Y.,  
RA Inamura S.-I., Kawana M., Masaki T., Kasanuki H., Sawamura T.,  
RA Matsuoka R.;  
RT "Oxidized LDL receptor gene (OLR1) is associated with the risk of  
RT myocardial infarction.";  
RL Biochem. Biophys. Res. Commun. 303:247-250(2003).  
RN [6]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
RX Millar D.S.;  
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Placenta;  
RX Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Siemsen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg H., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP FUNCTION.  
RX MEDLINE=21679116; PubMed=11821063;  
RA Hayaashida K., Kume N., Minami M., Kita T.;  
RT "Lectin-like oxidized LDL receptor-1 (LOX-1) supports adhesion of  
RT mononuclear leukocytes and a monocyte-like cell line THP-1 cells under  
RT static and flow conditions.";  
RL FEBS Lett. 511:133-138(2002).  
RN [9]  
RP FUNCTION, AND TISSUE SPECIFICITY.  
RX MEDLINE=22241698; PubMed=12354387;  
RA Delneste Y., Magistrelli G., Gauchat J.-P., Haeuw J.-P., Aubry J.-F.,  
RA Nakamura K., Kawakami-Honda N., Goetsch L., Sawamura T.,  
RA Bonnefoy J.-P., Jeannin P.;  
RT "Involvement of LOX-1 in dendritic cell-mediated antigen cross-  
RT presentation.";  
RL Immunity 17:353-362(2002).  
RN [10]  
RP INVOLVEMENT IN AD DISEASE.  
RX PubMed=12384789; DOI=10.1007/s00439-002-0802-7;  
RA Lueddecking-Zimmer E., DeRosky S.T., Chen Q., Barmada M.M.,  
RA Kamboh M.I.;  
RT "Investigation of oxidized LDL-receptor 1 (OLR1) as the candidate gene  
RT for Alzheimer's disease on chromosome 12.";  
RL Hum. Genet. 111:443-451(2002).  
RN [11]  
RP INDUCTION.  
RX MEDLINE=22760283; PubMed=12878212;  
RA Hu B., Li D., Sawamura T., Mehta J.L.;  
RT "Oxidized LDL through LOX-1 modulates LDL-receptor expression in human  
RT coronary artery endothelial cells.";  
RL Biochem. Biophys. Res. Commun. 307:1008-1012(2003).  
RN [12]  
RP INVOLVEMENT IN MYOCARDIAL INFARCTION.  
RX PubMed=12810610; DOI=10.1161/01.CIR.0000074207.85796.36;  
RA Chen Q., Reis S.E., Kammerer C., Craig W.Y., Lapierre S.E.,  
RA Zimmer E.L., McNamara D.M., Pauly D.F., Sharaf B., Holubkov R.,  
RA Bairey Merz C.N., Sopko G., Bontempo F., Kamboh M.I.;  
RT "Genetic variation in lectin-like oxidized low-density lipoprotein  
RT receptor 1 (LOX1) gene and the risk of coronary artery disease.";  
RL Circulation 107:3146-3151(2003).  
RN [13]  
RP INVOLVEMENT IN AD DISEASE.  
RX MEDLINE=22692680; PubMed=12807963;  
RA Lambert J.-C., Lueddecking-Zimmer E., Merrot S., Hayes A., Thaker U.,  
RA Desai P., Houzet A., Hermant X., Cottel D., Pritchard A., Iwatsubo T.,  
RA Pasquier F., Frigard B., Conneally P.M., Chartier-Harlin M.-C.,  
RA DeKeasy S.T., Lendon C., Mann D., Kamboh M.I., Anouyel P.;  
RT "Association of 3'-UTR polymorphisms of the oxidized LDL receptor 1  
RT (OLR1) gene with Alzheimer's disease.";  
RL J. Med. Genet. 40:424-430(2003).  
RN [14]  
RP HOMODIMERIZATION, INTERCHAIN DISULFIDE BOND, AND MUTAGENESIS OF  
RP CYS-140.  
RX PubMed=15000751; DOI=10.1089/104454904322759920;  
RA Xie Q., Matsunaga S., Nimi S., Ogawa S., Tokuyasu K., Sakakibara Y.,  
RA Machida S.;  
RT "Human lectin-like oxidized low-density lipoprotein receptor-1  
RT functions as a dimer in living cells.";

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
NCBI\_TaxID=5141;  
[1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=OR74N;  
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
RA Elkins T., Engels S., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,  
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
RA Kamal M., Kamysysselis M., Mauceli E., Bielek C., Rudd S., Frishman D.,  
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,  
RA Dasouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
RA Yarden O., Planann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
RA Natvig D.O., Alex L.A., Mannhaupt G., Erboile D.J., Freilich B.,  
RA Paulsen I., Sachs M.S., Lander E.S., Nussba C., Birren B.;  
RL "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
RT Nature 0:0-0(2003).  
CC - CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC EMBL: AABX01000470; EAA29020.1; -; Genomic\_DNA.  
DR InterPro: IPR003014; PAN.  
DR Pfam: PF00024; PAN; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 498 AA; 56281 MW; 468DFD06F34E22D7 CRC64;  
  
Query Match 38.6%; Score 51; DB 2; Length 498;  
Best Local Similarity 44.4%; Pred. No. 46;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 6 EVTAEDGTQRAEKCSKRPC 23  
|| :|:|:|:|:|  
Db 435 EVAVVEGVEXAEKCGKKC 452  
  
RESULT 13  
Q4QG99\_LEIMA  
ID Q4QG99\_LEIMA PRELIMINARY; PRT; 670 AA.  
AC Q4QG99;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=ImjF13.0550;  
OS Leishmania major.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
[1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Frieldin;  
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,  
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,  
RA Saunders D., Seeger K., Warren T., Apostolou Z., Bauser C., Beck A.,  
RA Blanchettin G., Borzym K., Bothe G., Bruschi C., Ciarloni L.,  
RA Duesterhoef A., Fuchs M., Gabel C., Goffeau A., Hilbert H.,  
RA Klages S., Kube M., Marra M., Masuy D., Mueller-Auer S., Pohl T.,  
RA Purnelle B., Reinhardt R., Rieger M., Robben J., Schaefer M.,  
RA Toso V., Volckaert G., Wamboldt R., Wedler H., Zimmermann M.,  
RA Rajandream M. and Barrell B.G.;  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL: CT005252; CAJ02599.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 670 AA; 71637 MW; 65EE3208EE5CE544 CRC64;  
  
Query Match 38.3%; Score 50.5; DB 2; Length 670;  
Best Local Similarity 41.4%; Pred. No. 76;  
Matches 12; Conservative 3; Mismatches 5; Indels 9; Gaps 1;

```

Qy 1 PLHN-----QEVTAEDGTORAEKCS 20
ID Q59F42 HUMAN PRELIMINARY; PRT; 676 AA.
Db 230 PLNLLIHLGLSDQVSLDRQTRAEKCS 258

RESULT 14
Q59F42 HUMAN
ID Q59F42 HUMAN PRELIMINARY; PRT; 676 AA.
AC Q59F42;
DC 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Zinc finger protein 7 (K0X 4, clone HF.16) variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Spleen;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209619; BAD92856.1; -; mRNA.
FT NON TER 1
SQ SEQUENCE 676 AA; 76582 MW; 0D0DB854951D2E72 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 676;
Best Local Similarity 47.6%; Pred. No. 92;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PLHNQEVTAEDGTORAEKCSK 21
Db 173 PLESQGESAEQMSQRCCECGK 193

RESULT 15
ZN7F HUMAN
ID ZN7F_HUMAN STANDARD; PRT; 686 AA.
AC P17097; P17015;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger protein 7 (Zinc finger protein K0X4) (Zinc finger protein
DE HF.16).
GN Name=ZN7F; Synonyms=K0X4;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Placenta;
RX MEDLINE=9016993; PubMed=2106481;
RA Lania L., Dotti E., Pannuti A., Pascucci A., Pengue G., Feliciello I.,
RA la Mantia G., Lanfrancone L., Pellicci P.-G.;
RT "cDNA isolation, expression analysis, and chromosomal localization of
RT two human zinc finger genes.";
RL Genomics 6:333-340(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC Tissue=Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 413-468.
RC Tissue=Lymphoid;
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RT cells.";
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Ubiquitously present in many human cell lines
CC of different embryological derivation.
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC family.
CC -!- SIMILARITY: Contains 15 C2H2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC This Swiss-Prot entry is a copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M29580; AAA61313.1; -; mRNA.
DR EMBL; BC058923; AAH58923.1; -; mRNA.
DR EMBL; X52335; CAA36561.1; -; mRNA.
DR FIR; A34612; A34612.
DR HSSP; P08047; LSP2.
DR SMR; P17097; 311-413.
DR TRANSFAC; T04979; -.
DR Ensembl; ENSG00000147789; Homo sapiens.
DR HGNC; HGNC:13139; ZNF7.
DR MIM; 194531; -.
DR GO; GO:0003677; F:DNA binding; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 14.
DR PRINTS; PR00048; ZINGFINGER.
DR ProDom; PD000003; Znf_C2H2; 14.
DR SMART; SM00349; KRAB_1.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 14.
KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT DOMAIN 4 76 KRAB.
FT ZN_FING 223 245 C2H2-type 1.
FT ZN_FING 250 272 C2H2-type 2.
FT ZN_FING 278 300 C2H2-type 3.
FT ZN_FING 306 328 C2H2-type 4.
FT ZN_FING 334 356 C2H2-type 5.
FT ZN_FING 362 384 C2H2-type 6.
FT ZN_FING 413 435 C2H2-type 7.
FT ZN_FING 441 463 C2H2-type 8.
FT ZN_FING 469 491 C2H2-type 9.
FT ZN_FING 497 519 C2H2-type 10.

```

AL QOCNEL;

DR GO: 0004872; F: receptor activity; IEA.  
 DR GO: 0004714; F: transmembrane receptor protein tyrosine kin. . . ; IEA.  
 DR GO: 0006468; P: protein amino acid phosphorylation; IEA.  
 DR GO: 0007169; P: transmembrane receptor protein tyrosine kin. . . ; IEA.  
 DR InterPro: IPR006211; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 KW Receptor.  
 FT NON TER 1 1  
 FT NON TER 74 74  
 SQ SEQUENCE 74 AA; 8155 MW; B3BEF7953BBCOAE CRC64;  
 Query Match 37.1%; Score 49; DB 2; Length 74;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 10 EDGTQRAEKCSPC 23  
 Db 31 EDGVRCKCKEGPC 44  
 RESULT 19  
 Q9BG67\_RABIT PRELIMINARY; PRT; 116 AA.  
 ID Q9BG67; AC Q9BG67;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Receptor tyrosine kinase ErbB1 (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;  
 OC Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,  
 RA Tetens F., Fischer B.,  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF333177; AAK14370.1; -; mRNA.  
 DR HSP; Q9H2C9; IMOX.  
 DR SMR; Q9BG67; 1-113.  
 DR GO: 0016021; C: integral to membrane; IEA.  
 DR GO: 0016020; C: membrane; IEA.  
 DR GO: 0005524; F: ATP binding; IEA.  
 DR GO: 0016301; F: kinase activity; IEA.  
 DR GO: 0004872; F: receptor activity; IEA.  
 DR GO: 0004714; F: transmembrane receptor protein tyrosine kin. . . ; IEA.  
 DR GO: 0006468; P: protein amino acid phosphorylation; IEA.  
 DR GO: 0007169; P: transmembrane receptor protein tyrosine kin. . . ; IEA.  
 DR InterPro: IPR006211; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 KW Kinase; Receptor.  
 FT NON TER 1 1  
 FT NON TER 116 116  
 SQ SEQUENCE 116 AA; 12624 MW; 6B61F79C0FF2322F CRC64;  
 Query Match 37.1%; Score 49; DB 2; Length 116;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 10 EDGTQRAEKCSPC 23  
 Db 42 EDGVRCKCKEGPC 55  
 RESULT 20  
 VXIS\_BPSF5  
 ID VXIS\_BPSF5 STANDARD; PRT; 147 AA.  
 AC Q2010;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Excisionase.

GN Name=XIS; Synonyms=27;  
 OS Bacteriophage SfV (Shigella flexneri bacteriophage V).  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.  
 OX NCBI\_TaxID=55884;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=97449297; PubMed=9305766; DOI=10.1016/S0378-1119(97)00144-3;  
 RA Huan P.T., Whittle B.L., Bastin D.A., Lindberg A.A., Verma N.K.;  
 RT "Shigella flexneri type-specific antigen V: cloning, sequencing and  
 RT characterization of the glucosyl transferase gene of temperate  
 RT bacteriophage SfV.";  
 RL Gene 195:207-216 (1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX MEDLINE=21886391; PubMed=11889106;  
 DOI=10.1128/JB.184.7.1974-1987.2002;  
 RA Allison G.E., Angeles D., Tran-Dinh N., Verma N.K.;  
 RT "Complete genomic sequence of SfV, a serotype-converting temperate  
 RT bacteriophage of Shigella flexneri.";  
 RL J. Bacteriol. 184:1974-1987 (2002).  
 CC -|- FUNCTION: Excisionase and integrase are necessary for the excision  
 CC of prophage from the host genome by site-specific recombination  
 CC (by similarity).  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL: U82619; AAB72136.1; -; Genomic DNA.  
 KW DNA recombination; DNA-binding.  
 SQ SEQUENCE 147 AA; 16524 MW; 37A541BBBABC08F7 CRC64;  
 Query Match 37.1%; Score 49; DB 1; Length 147;  
 Best Local Similarity 47.4%; Pred. No. 26;  
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 PLHNQEVTAEDGTQRAEKC 19  
 Db 90 PLHTVQVSAGDGITEERKC 108  
 RESULT 21  
 Q68GS5\_HUMAN  
 ID Q68GS5\_HUMAN PRELIMINARY; PRT; 405 AA.  
 AC Q68GS5;  
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE Epidermal growth factor receptor short isoform.  
 GN Name=EGFR;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Xu L., Hong A., He X.;  
 RT "Cloning of the cDNA for a short EGF receptor from human placenta.";  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY698024; AAT97979.1; -; mRNA.  
 KW Receptor.  
 SQ SEQUENCE 405 AA; 44664 MW; F5DEB31787EF1822 CRC64;  
 Query Match 37.1%; Score 49; DB 2; Length 405;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 10 EDGTQRAEKCSPC 23  
 Db 320 EDGVRCKCKEGPC 333

```
RT      "None Title.";
RL      Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB209442; BAD92679.1; -; mRNA.
DR      SMR; Q59FL8; 25-509.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0005489; F:electron transporter activity; IEA.
DR      GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR      GO; GO:0005506; F:iron ion binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0004872; F:receptor activity; IEA.
DR      GO; GO:0006119; F:electron transport; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR      InterPro; IPR001450; 4Fe4S_Fe_S_bd.
DR      InterPro; IPR000494; EGFR_L.
DR      InterPro; IPR006211; Furin-like.
DR      InterPro; IPR006212; Furin repeat.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_pkinase.
DR      InterPro; IPR001245; Tyr_pkinase.
DR      InterPro; IPR008266; Tyr_pkinase_AS.
DR      Pfam; PF00757; Furin-like; 1.
DR      Pfam; PF01030; Recep_L_domain; 1.
DR      PRINTS; PR00353; 4FE4SFRDOXIN.
DR      PRINTS; PR00109; TYRKINASE.
DR      SMART; SM00261; FU; 3.
DR      SMART; SM00220; S_TKc; 1.
DR      SMART; SM00219; TyrKc; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW      ATP-binding; Receptor.
FT      NON_TER 1
SQ      SEQUENCE 1081 AA; 119952 MW; C3BAF46C211A86DC CRC64;

Query Match      37.1%; Score 49; DB 2; Length 1081;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      10 EDGTORAEKCKSPC 23
      ||| : : : |||
DB      191 EDGVRKCKCEGFC 204

RESULT 24
Q504U8 HUMAN
ID      Q504U8_HUMAN PRELIMINARY; PRT; 1091 AA.
AC      Q504U8;
DT      13-SEP-2005 (TrEMBLrel. 31, Created)
DT      13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      EGFR protein.
GN      Name=EGFR;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC      Homo.
OC      NCBI_TaxID=9606;
OX      [1]
RN      NUCLEOTIDE SEQUENCE.
RP      TISSUE=Placenta;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Placenta;  
RG NIH MGC Project;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC094761; AAH94761.1; -; mRNA.  
DR SMR; Q504U8; 26-593.  
DR InterPro; IPR001450; 4Fe4S\_Fe\_S\_bd.  
DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR006211; Furin-like.  
DR InterPro; IPR006212; Furin repeat.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR002290; Ser\_Thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
DR Pfam; PF01030; Recep\_L domain; 2.  
DR PRINTS; PR00757; Furin-like; 1.  
DR PRINTS; PR00353; 4Fe4S\_FeDOXIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; FU; 3.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
SQ SEQUENCE 1091 AA; 258062C735D62AD1 CRC64;  
Query Match 37.1%; Score 49; DB 2; Length 1091;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 10 EDGTQRAEKCKGPC 23  
||| : : : ||  
DB 275 EDGVRKCKCEGPC 288

RESULT 25  
Q8MIL8\_PIG  
ID Q8MIL8\_PIG PRELIMINARY; PRT; 1209 AA.  
AC Q8MIL8; 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Epidermal growth factor receptor.  
OS Sus scrofa (pig)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
OC Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Kim J.G., Vallet J.L., Nonnenan D., Christenson R.K.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY117054; AAH74742.1; -; mRNA.  
DR HSSP; Q9H2C9; 1M17.  
DR SMR; Q8MIL8; 25-525, 27-636.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.

DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR006211; Furin-like.  
DR InterPro; IPR006212; Furin repeat.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
DR Pfam; PF00757; Furin-like; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR SMART; SM00261; FU; 5.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Receptor.  
SQ SEQUENCE 1209 AA; 133531 MW; 268E3FB11E36F90F CRC64;  
Query Match 37.1%; Score 49; DB 2; Length 1209;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 10 EDGTQRAEKCKGPC 23  
||| : : : ||  
DB 320 EDGVRKCKCEGPC 333

RESULT 26  
EGFR\_HUMAN  
ID EGFR\_HUMAN STANDARD; PRT; 1210 AA.  
AC P00533; Q00688; Q00732; P06268; Q14225; Q92795; Q9BZS2; Q9GZX1;  
AC Q9H2C9; Q9H3C9; Q9UMD7; Q9UMD8; Q9UMG5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
DE tyrosine-protein kinase ErbB-1).  
GN Name=EGFR; Synonyms=ERBB1;  
OS Homo sapiens (human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
RX MEDLINE=84219729; PubMed=6328312;  
RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
RA Lee J., Varden Y., Libermann T.A., Schlessinger J., Downward J.,  
RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
RT expression of the amplified gene in A431 epidermoid carcinoma cells.";  
RL Nature 309:418-425(1984).  
RN [2]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Placenta;  
RX MEDLINE=95382957; PubMed=7654368; DOI=10.1262/jrd.41.149;  
RA Ilekis J.V., Stark B.C., Scoccia B.;  
RT "Possible role of variant RNA transcripts in the regulation of  
RT epidermal growth factor receptor expression in human placenta.";  
RL Mol. Reprod. Dev. 41:149-156(1995).  
RN [3]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Placenta;  
RX MEDLINE=97078686; PubMed=8918811; DOI=10.1093/nar/24.20.4050;  
RA Reiter J.L., Mailhe N.J.;  
RT "A 1.8 kb alternative transcript from the human epidermal growth  
RT factor receptor gene encodes a truncated form of the receptor.";  
RL Nucleic Acids Res. 24:4050-4056(1996).  
RN [4]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Placenta;  
RX MEDLINE=97256547; PubMed=9103388; DOI=10.1006/gyno.1996.4526;  
RA Ilekis J.V., Gariti J., Niederberger C., Scoccia B.;

RT "Expression of a truncated epidermal growth factor receptor-like  
RT protein (TEGFR) in ovarian cancer.";  
RN Gynecol. Oncol. 65:36-41(1997).  
RN [5]  
RC TISSUE=Placenta;  
RC NUCLEOTIDE SEQUENCE (ISOFORMS 3 AND 4).  
RC  
RX MEDLINE=21100872; PubMed=11161793; DOI=10.1006/geno.2000.6341;  
RX Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
RA Maihle N.J.;  
RT "Comparative genomic sequence analysis and isolation of human and  
RT mouse alternative EGFR transcripts encoding truncated receptor  
RT isoforms.";  
RT Genomics 71:11-20(2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS GLN-98; ARG-266;  
RP LYS-521; ILE-674; GLY-962 AND PRO-988.  
RA Livingstone R.J., Rieder M.J., Chung M.-W., Ritchie T.K., Olson A.N.,  
RA Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S.,  
RA Sherwood J.K., Sherwood A.M., Leithauser B.J., Nickerson D.A.;  
RT "NIHNS SNPs, environmental genome project. NIHNS ES15478, Department  
RT of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu>).";  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RN NUCLEOTIDE SEQUENCE OF 575-687.  
RA Reiter J.L., Threadgill D.W., Danielsen A.J., Schehl C.M.,  
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
RA Maihle N.J.;  
RT "Human and mouse alternative EGFR transcripts encoding only the  
RT extracellular domain of the receptor.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP NUCLEOTIDE SEQUENCE OF 713-924.  
RX MEDLINE=84196372; PubMed=6326261;  
RX Lin C.R., Chen W.S., Krutiger W.; Stolarsky L.S., Weber W., Evans R.M.,  
RA Verma I.M., Gill G.N., Rosenfeld M.G.;  
RT "Expression cloning of human EGF receptor complementary DNA: gene  
RT amplification and three related messenger RNA products in A431  
RT cells.";  
RL Science 224:843-848(1984).  
RN [9]  
RN NUCLEOTIDE SEQUENCE OF 150-962.  
RX MEDLINE=84245835; PubMed=6330563;  
RX Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
RA Roe B.A., Merlino G.T., Pastan I.;  
RT "Human epidermal growth factor receptor cDNA is homologous to a  
RT variety of RNAs overproduced in A431 carcinoma cells.";  
RL Nature 309:806-810(1984).  
RN [10]  
RP NUCLEOTIDE SEQUENCE OF 1028-1210.  
RX MEDLINE=85046483; PubMed=6093780;  
RX Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,  
RA O'Malley B.W.;  
RT "Isolation of an evolutionarily conserved epidermal growth factor  
RT receptor cDNA from human A431 carcinoma cells.";  
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
RN [11]  
RP NUCLEOTIDE SEQUENCE OF 1-29.  
RX MEDLINE=88217333; PubMed=3329716;  
RX Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
RA Waterfield M.D.;  
RT "The human EGF receptor gene: structure of the 110 kb locus and  
RT identification of sequences regulating its transcription.";  
RL Oncogene Res. 1:375-396(1987).  
RN [12]  
RP NUCLEOTIDE SEQUENCE OF 1-29.  
RX MEDLINE=91107677; PubMed=1988448;  
RX Haley J.D., Waterfield M.D.;  
RT "Contributory effects of de novo transcription and premature  
RT transcript termination in the regulation of human epidermal growth  
RT factor receptor proto-oncogene RNA synthesis.";  
RL J. Biol. Chem. 266:1746-1753(1991).

RN [13]  
RP NUCLEOTIDE SEQUENCE OF 1-29.  
RX MEDLINE=85270438; PubMed=2991899;  
RX Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
RT "Characterization and sequence of the promoter region of the human  
RT epidermal growth factor receptor gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
RN [14]  
RP NUCLEOTIDE SEQUENCE OF 25-49.  
RX MEDLINE=84172183; PubMed=6324343;  
RX Weber W., Gill G.N., Spiess J.;  
RT "Production of an epidermal growth factor receptor-related protein.";  
RL Science 224:294-297(1984).  
RN [15]  
RP PROTEIN SEQUENCE OF 540.  
RA Kohda D.;  
RL Submitted (SEP-1997) to Swiss-Prot.  
RN [16]  
RP PROTEIN SEQUENCE OF 687-705; 986-998; 1000-1023; 1026-1030 AND  
RP 1068-1077, AND PHOSPHORYLATION SITES THR-693; SER-695; SER-1070 AND  
RP SER-1071.  
RX MEDLINE=88330814; PubMed=3138233;  
RX Heisermann G.J., Gill G.N.;  
RA "Epidermal growth factor receptor threonine and serine residues  
RT phosphorylated in vivo.";  
RL J. Biol. Chem. 263:13152-13158(1988).  
RN [17]  
RP PROTEIN SEQUENCE OF 25-39.  
RX PubMed=15340161; DOI=10.1110/ps.04682504;  
RX Zhang Z., Hensel W.J.;  
RT "Signal peptide prediction based on analysis of experimentally  
RT verified cleavage sites.";  
RL Protein Sci. 13:2819-2824(2004).  
RN [18]  
RP PROTEIN SEQUENCE OF 740-744 AND 746-747.  
RX MEDLINE=85182850; PubMed=2985580;  
RX Russo M.W., Lukas T.J., Cohen S., Staros J.V.;  
RT "Identification of residues in the nucleotide binding site of the  
RT epidermal growth factor receptor/kinase.";  
RL J. Biol. Chem. 260:5205-5208(1985).  
RN [19]  
RP RECEPTOR ACTIVITY.  
RX MEDLINE=84191554; PubMed=6325948;  
RX Mroczkowski B., Mosig G., Cohen S.;  
RT "ATP-stimulated interaction between epidermal growth factor receptor  
RT and supercoiled DNA.";  
RL Nature 309:270-273(1984).  
RN [20]  
RP LIGAND BINDING.  
RX MEDLINE=90003233; PubMed=2790960; DOI=10.1016/0092-8674(89)90867-2;  
RX Chen W.S., Lazar C.S., Lund K.A., Welsh J.B., Chang C.P., Walton G.M.,  
RA Der C.J., Wiley H.S., Gill G.N., Rosenfeld M.G.;  
RT "Functional independence of the epidermal growth factor receptor from  
RT a domain required for ligand-induced internalization and calcium  
RT regulation.";  
RL Cell 59:33-43(1989).  
RN [21]  
RP PHOSPHORYLATION.  
RX MEDLINE=89278137; PubMed=2543678;  
RX Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
RA Hawk R., Givol D., Ullrich A., Schlessinger J.;  
RT "All autophosphorylation sites of epidermal growth factor (EGF)  
RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
RT Identification of a novel site in EGF receptor.";  
RL J. Biol. Chem. 264:10667-10671(1989).  
RN [22]  
RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
RP ASN-528.  
RX MEDLINE=96398132; PubMed=8962717;  
RX Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
RT "Analysis of the glycosylation patterns of the extracellular domain of  
RT the epidermal growth factor receptor expressed in Chinese hamster  
RT ovary fibroblasts.";



*Eutelleostomus* (Indonesian Coelacanth).

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:



```
RA Pellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003)).
CC -I- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DDSI third party annotation (TPA) entry.
DR EMBL; BK003679; DAA02377.1; -; Genomic DNA.
SQ SEQUENCE 150 AA; 16067 MW; 5657862531B2DD35 CRC64;

Query Match 36.4%; Score 48; DB 2; Length 150;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 5 QEVTAEDGTQRAEKCSPCA 24
DB 25 QCMVEEMGTSPAERCSQPA 44

RESULT 33
Q8XDI3 EC057 PRELIMINARY; PRT; 214 AA.
AC O8XDI3; 07AEV8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE , complete genome (Hypothetical protein ECs1696).
GN OrderedLocusNames=ECs1696, z1963;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074933; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Gottlieb E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouais K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005174; AAC56051.1; -; Genomic DNA.
DR EMBL; BA000007; BAB35119.1; -; Genomic DNA.
DR PIR; G85698; G85698.
DR PIR; H90840; H90840.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000051; SAM bind.
KW Complete proteome; Hypothetical protein; Methyltransferase;
KW Transferase.
SQ SEQUENCE 214 AA; 23967 MW; 229CAF9C77704BE0 CRC64;

Query Match 36.4%; Score 48; DB 2; Length 214;
Best Local Similarity 45.5%; Pred. No. 55;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 HNOEVTAEADGTQRAEKCSPCA 24
DB 20 HRTEKTPDHWQRAEKMAENCA 41

RESULT 34
Q83LE8 SHIFL PRELIMINARY; PRT; 269 AA.
AC Q83LE8; Q7C122;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Sl277, SF1191;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Barling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; AAN42805.1; -; Genomic DNA.
DR EMBL; AE015982; AAP16693.1; -; Genomic DNA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000051; SAM bind.
KW Complete proteome; Hypothetical protein; Methyltransferase;
KW Transferase.
SQ SEQUENCE 269 AA; 30722 MW; D5352DAAFC1D745C CRC64;

Query Match 36.4%; Score 48; DB 2; Length 269;
Best Local Similarity 45.5%; Pred. No. 70;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 HNOEVTAEADGTQRAEKCSPCA 24
DB 20 HRTEKTPDHWQRAEKMAENCA 41

RESULT 35
Q87503 ECOLI PRELIMINARY; PRT; 270 AA.
AC Q87503;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orf2.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CFT073;
RX MEDLINE=97342757; PubMed=9199454;
```

RESULT 37  
Q9ESE0 RAT  
ID Q9ESE0 RAT PRELIMINARY; PRT; 478 AA.  
AC Q9ESE0;  
DT 01-MAR-2001 (TREMBLrel. 16. Created)

```

RESULT 38
ID Q6NKT4_ ARATH PRELIMINARY; PRT; 632 AA.
Q6NKT4_ ARATH
OC GNKNT4;
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE At4g14650.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes; Magnoliophyta; eudicotyledons; core eudicotyledons II; Brassicales; Brassicaceae; Arabidopsi.
OX euroids II;
NCBI_TaxID=3702;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Eckert J.R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DDJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Eckert J.R.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DDJ databases.
[3]
RP "Arabidopsis ORF clones."
RL Submitted (FEB-2005) to the EMBL/GenBank/DDJ databases.
DR EMBL; BT012609; AAT06428.1; -; mRNA.
DR EMBL; BT020594; AAW80867.1; -; mRNA.
SQ SEQUENCE 632 AA; 70003 MW; 08F4F98935AA639B CRC64;

```

```

Query Match      36.4%; Score 48; DB 2; Length 632;
Best Local Similarity 39.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 LHNQEVTAEDGTQRAEKCKPCA 24
Db 393 VHDEVTDKGADKAISNKPSCS 415

RESULT 39
QSERV6_MOUSE
ID QSERV6_MOUSE PRELIMINARY; PRT; 643 AA.
AC QSERV6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Epidermal growth factor receptor isoform 2.
GN Name=Egfr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C3H/101;
RX MEDLINE=21100872; PubMed=11161793; DOI=10.1006/geno.2000.6341;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,
RA Malhale N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative EGFR transcripts encoding truncated receptor
RT isoforms.";
RL Genomics 71.1-20(2001).
DR EMBL; AF275366; AAG28046.1; -; Genomic_DNA.
DR EMBL; AF275364; AAG28046.1; JOINED; Genomic_DNA.
DR EMBL; AF275365; AAG28046.1; JOINED; Genomic_DNA.
DR HSSP; O9H2C9; IMOX.
DR SMR; QSERV6; 27-631.
DR MGI; MGI:95294; Egfr.
DR GO; GO:001623; C:basolateral plasma membrane; IDA.
DR GO; GO:0030139; C:endocytic vesicle; IDA.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005622; C:intracellular; IDA.
DR GO; GO:0016301; F:kinase activity; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0004871; F:signal transducer activity; IDA.
DR GO; GO:0050730; P:regulation of peptidyl-tyrosine phosphoryla. .; IMP.
DR GO; GO:0007165; P:signal transduction; IDA.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin-repeat.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00353; 4FE4SFERDOXIN.
DR SMART; SM00261; FU; 3.
KW Receptor.
SQ SEQUENCE 643 AA; 71476 MW; DEF22002C84911B1 CRC64;

Query Match      36.4%; Score 48; DB 2; Length 643;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 10 EDCQRAEKCKPCA 23
Db 320 EDGIRCKCKDGPC 333

RESULT 40
O23321_ARATH
ID O23321_ARATH PRELIMINARY; PRT; 644 AA.

```

```

O23321;
AC 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein dl3365c (Hypothetical protein AT4g14650).
GN Name=dl3365c; Synonym=AT4g14650;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bevan M., Stiekema W., Murphy G., Wamburt R., Pohl T., Terry N.,
RA Kreis M., Kavanagh T., Estian K.D., Rieger M., James R.,
RA Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA Jones J., Palme K., Ansoorge W., Delsey M., Bancroft I., Mewes H.W.,
RA Schueller C., Chalwatizis N.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z97336; CAB10244.1; -; Genomic_DNA.
DR EMBL; AL161539; CAB78507.1; -; Genomic_DNA.
DR PIR; B71409; B71409.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 71277 MW; 8057168456922960 CRC64;

Query Match      36.4%; Score 48; DB 2; Length 644;
Best Local Similarity 39.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 LHNQEVTAEDGTQRAEKCKPCA 24
Db 405 VHDEVTDKGADKAISNKPSCS 427

Search completed: December 3, 2005, 07:48:42
Job time : 167.273 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM' protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 22:22:28 ; Search time 2200.91 Seconds  
(without alignments)  
619.854 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLHQEVTAEDGTQAEKCKPCA 24

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFPO\_spool/US09632036/runat\_02122005\_104140\_25811/app.query.fasta\_1.398  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09632036 @CGN 1 1 3563 @runat\_02122005\_104140\_25811 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	471	6	AX379026 Sequence
2	128	97.0	1219	6	CS085264 Sequence
3	128	97.0	1316	8	AF177761 Homo sapi

CS075259 Sequence	6	CS075259	1320	97.0	128
CS085263 Sequence	6	CS085263	1459	97.0	128
CS085262 Sequence	6	CS085262	1699	97.0	128
AR181000 Sequence	6	AR181000	1872	97.0	128
CS085261 Sequence	6	CS085261	1939	97.0	128
CS075237 Sequence	6	CS075237	1944	97.0	128
CS020247 Sequence	6	CS020247	2028	97.0	128
CS020248 Sequence	6	CS020248	2028	97.0	128
CS075249 Sequence	6	CS075249	2061	97.0	128
CS085268 Sequence	6	CS085268	2086	97.0	128
CS085269 Sequence	6	CS085269	2086	97.0	128
CS085270 Sequence	6	CS085270	2086	97.0	128
CS085271 Sequence	6	CS085271	2086	97.0	128
CS075235 Sequence	6	CS075235	2320	97.0	128
AR082744 Sequence	6	AR082744	2385	97.0	128
AR099963 Sequence	6	AR099963	2385	97.0	128
AR143949 Sequence	6	AR143949	2385	97.0	128
CS075247 Sequence	6	CS075247	2437	97.0	128
BD222937 Heteromin	6	BD222937	2871	97.0	128
AX023363 Sequence	6	AX023363	2871	97.0	128
AX051114 Sequence	6	AX051114	3678	97.0	128
AR034479 Sequence	6	AR034479	3768	97.0	128
BD224136 Novel met	6	BD224136	3768	97.0	128
BD267514 HER-2/neu	6	BD267514	3768	97.0	128
CQ815384 Sequence	6	CQ815384	3768	97.0	128
CQ841302 Sequence	6	CQ841302	3768	97.0	128
CQ841341 Sequence	6	CQ841341	3768	97.0	128
CQ841343 Sequence	6	CQ841343	3768	97.0	128
CQ841344 Sequence	6	CQ841344	3768	97.0	128
CS007994 Sequence	6	CS007994	3768	97.0	128
CS020239 Sequence	6	CS020239	3768	97.0	128
CS104222 Sequence	6	CS104222	3768	97.0	128
AR409603 Sequence	6	AR409603	3768	97.0	128
AX060704 Sequence	6	AX060704	3768	97.0	128
AX021817 Sequence	6	AX021817	3768	97.0	128
AX380923 Sequence	6	AX380923	3768	97.0	128
AX384604 Sequence	6	AX384604	3768	97.0	128
AX465456 Sequence	6	AX465456	3768	97.0	128
AX467229 Sequence	6	AX467229	3768	97.0	128
AX481438 Sequence	6	AX481438	3768	97.0	128
AY888050 Synthetic	11	AY888050	3768	97.0	128
AY888051 Synthetic	11	AY888051	3768	97.0	128
AY888102 Synthetic	11	AY888102	3768	97.0	128
AY888103 Synthetic	11	AY888103	3768	97.0	128
AY888104 Synthetic	11	AY888104	3768	97.0	128
CS020249 Sequence	6	CS020249	3778	97.0	128
AR080259 Sequence	6	AR080259	4473	97.0	128
AR167390 Sequence	6	AR167390	4473	97.0	128
AR392088 Sequence	6	AR392088	4473	97.0	128
AR534814 Sequence	6	AR534814	4473	97.0	128
X03363 Human c-erb	8	HSERB2R	4473	97.0	128
CQ723568 Sequence	6	CQ723568	4483	97.0	128
CQ834023 Sequence	6	CQ834023	4530	97.0	128
CQ873909 Sequence	6	CQ873909	4530	97.0	128
CQ977093 Sequence	6	CQ977093	4530	97.0	128
CS052366 Sequence	6	CS052366	4530	97.0	128
CS063102 Sequence	6	CS063102	4530	97.0	128
CS080883 Sequence	6	CS080883	4530	97.0	128
CS101342 Sequence	6	CS101342	4530	97.0	128
CS103534 Sequence	6	CS103534	4530	97.0	128
I21124 Sequence 9	6	I21124	4530	97.0	128
I59745 Sequence 9	6	I59745	4530	97.0	128
AR202597 Sequence	6	AR202597	4530	97.0	128
AR283481 Sequence	6	AR283481	4530	97.0	128
AR344811 Sequence	6	AR344811	4530	97.0	128
AX282577 Sequence	6	AX282577	4530	97.0	128
AX587649 Sequence	6	AX587649	4530	97.0	128
AX644071 Sequence	6	AX644071	4530	97.0	128
AX771418 Sequence	6	AX771418	4530	97.0	128
AX938565 Sequence	6	AX938565	4530	97.0	128
BD005474 Cellular	6	BD005474	4530	97.0	128
M11730 Human tyros	8	HUMHER2A	4530	97.0	128
BC080193 Homo sapi	8	BC080193	4536	97.0	128

77	128	97.0	4647	6	CS075243	CS075243	Sequence
78	128	97.0	4932	8	CQ953478	CQ953478	Sequence
79	128	97.0	4932	8	AK131568	AK131568	Homo sapi
80	128	97.0	5661	6	CQ975755	CQ975755	Sequence
81	128	97.0	9274	6	AR409602	AR409602	Sequence
82	128	97.0	9274	6	AX060703	AX060703	Sequence
83	128	97.0	30837	8	AY208911	AY208911	Homo sapi
84	128	97.0	89963	8	CR626867	CR626867	Human DNA
85	128	97.0	151284	8	CR626880	CR626880	Human DNA
86	128	97.0	154278	14	AC025531	AC025531	Homo sapi
87	128	97.0	157216	8	AC087491	AC087491	Homo sapi
88	128	97.0	161815	8	AC079199	AC079199	Homo sapi
89	128	97.0	168585	8	AC040933	AC040933	Homo sapi
90	128	97.0	198008	14	AC142197	AC142197	Homo sapi
91	121	91.7	3780	4	AB008451	AB008451	Canis fam
92	121	91.7	3783	4	AY702651	AY702651	Felis cat
93	114	86.4	641	10	BV070407	BV070407	S212P6048
94	114	86.4	2083	6	CS085260	CS085260	Sequence
95	114	86.4	2086	6	CS085272	CS085272	Sequence
96	114	86.4	2783	6	AX380942	AX380942	Sequence
97	114	86.4	2781	6	AX380944	AX380944	Sequence
98	114	86.4	3771	6	BD267516	BD267516	HER-2/neu
99	114	86.4	3771	6	AX189662	AX189662	Sequence
100	114	86.4	3771	6	AX380925	AX380925	Sequence



```
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21-q22"
1..1316
/gene="HER-2"
/notes="formerly neu; synonym: c-erbB-2; developmentally
expressed oncogene"
13..1272
/gene="HER-2"
/functions="binds p185HER-2 receptor"
/functions="inhibits HER-2-mediated tumorigenic growth"
/notes="approximates 68 kDa secreted autoinhibitor
glycoprotein; alternatively spliced"
/codon_start=1
/product="heretatin"
/protein_id="AADS6009.2"
/db_xref="GI:10181233"
/translators="MEALCRLWGLLALLPPGAASQVCTGDMKRLPASPETHLD
MLRHLVGGCQVQGNLELTYPNTNASLFQDIQEVQYVLIAHNQVRQVPLQRLRIV
RGQLFEDNVALAVLQNGDPLNNTPTVTSRACHPCSPCKGSRGWSGESSDCOSLRT
LCYQDTILWKDIIPHKNQALTLIDTNRSPACHPCSPCKGSRGWSGESSDCOSLRT
VCAGGARCPPLPTDCHEQCAAGCTGPKHSDCLACHPHNSGICELHCPALVTYNT
DTESMPNPGRTYFGASCVTACPYNTLSTVGSCTLVCPLNQEVTAEDGTQRCBKC
SKPCARHTSLPRPAVPVPLRMQGPAPHPVLSFLRPSWDLVSAFYSLPLPLSPTS
VPIPSVSVGRGPDPAHVAVDLSRYEG"

ORIGIN
Alignment Scores:
Pred. No.: 1.26e-10 Length: 1316
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 8 Gaps: 0

US-09-632-036F-6 (1-24) x AF177761 (1-1316)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 958 CCCCTGCACACCAAGAGGTGACAGAGAGTGGAAACACACGCGTGTGAGAGTGCAGC 1017
|||||

QY 21 LysProCysAla 24
|||||
Db 1018 AAGCCCTGTGCC 1029
|||||

RESULT 4
CS075259 1320 bp DNA linear PAT 05-MAY-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1
REFERENCE
AUTHORS
Shemesh,R., Oren,A., Rotman,G., Sela-Tavor,O., Walach,S.,
Samesch-Greenwald,S., Beiman,M., Eshel,D. and Savitsky,K.
TITLE
Polynucleotides encoding novel ErbB-2 polypeptides and kits and
methods using same
JOURNAL
Patent: WO 200503133-A 25 14-APR-2005;
Compugen Ltd. (IL)
FEATURES
Location/Qualifiers
1..1320
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
```

```
Pred. No.: 1.27e-10 Length: 1320
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CS075259 (1-1320)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 1096 CCCCTGCACACCAAGAGGTGACAGAGAGTGGAAACACACGCGTGTGAGAGTGCAGC 1155
|||||

QY 21 LysProCysAla 24
|||||
Db 1156 AAGCCCTGTGCC 1167
|||||

RESULT 5
CS085263 1459 bp DNA linear PAT 25-MAY-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CS085263
Sequence 5 from Patent WO2005039618.
CS085263
CS085263.1 GI:66711220
unidentified
unclassified.
REFERENCE
AUTHORS
Amici,A., Cavallo,F., Forni,G. and Marchini,C.
TITLE
P185 sp neu /sp -encoding DNA and therapeutic uses thereof
JOURNAL
Patent: WO 2005039618-A 5 06-MAY-2005;
Indena S.p.A. (IT)
FEATURES
Location/Qualifiers
1..1459
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 1.41e-10 Length: 1459
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CS085263 (1-1459)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 350 CCCCTGCACACCAAGAGGTGACAGAGAGTGGAAACACACGCGTGTGAGAGTGCAGC 409
|||||

QY 21 LysProCysAla 24
|||||
Db 410 AAGCCCTGTGCC 421
|||||

RESULT 6
CS085262 1699 bp DNA linear PAT 25-MAY-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CS085262
Sequence 4 from Patent WO2005039618.
CS085262
CS085262.1 GI:66711219
unidentified
unclassified.
REFERENCE
AUTHORS
Amici,A., Cavallo,F., Forni,G. and Marchini,C.
TITLE
P185 sp neu /sp -encoding DNA and therapeutic uses thereof
JOURNAL
Patent: WO 2005039618-A 4 06-MAY-2005;
Indena S.p.A. (IT)
FEATURES
Location/Qualifiers
1..1699
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
```



```
RESULT 10
CS020247
LOCUS      2028 bp      DNA      linear      PAT 23-FEB-2005
DEFINITION Sequence 9 from Patent WO2005012527.
ACCESSION CS020247
VERSION    CS020247.1 GI:60220905
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   other sequences; artificial sequences.

REFERENCE
1 Gallo,P.I., Monaci,P.I. and Nuzzo,M.I.
  Synthetic gene encoding human epidermal growth factor 2/neu antigen
  and uses thereof
JOURNAL    Patent: WO 2005012527-A 9 10-FEB-2005;
            Istituto di Ricerche di Biologia Molecolare P. Angeletti S.P.A.
            (IT)

FEATURES
source     Location/Qualifiers
1 .2028
   /organism="synthetic construct"
   /mol_type="unassigned DNA"
   /db_xref="taxon:32630"
   /note="HER2ECDTwt"

ORIGIN
Alignment Scores:
Pred. No.: 2e-10      Length: 2028
Score: 128.00        Matches: 23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%    Mismatches: 1
Query Match: 96.97%      Indels: 0
DB: 6                Gaps: 0

US-09-632-036F-6 (1-24) x CS020247 (1-2028)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGCACAACCAAGAGGTGACCGCGAGGACGCCAGCCAGCGCTGCGAGAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 11
CS020248
LOCUS      2028 bp      DNA      linear      PAT 23-FEB-2005
DEFINITION Sequence 10 from Patent WO2005012527.
ACCESSION CS020248
VERSION    CS020248.1 GI:60220906
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   other sequences; artificial sequences.

REFERENCE
1 Gallo,P.I., Monaci,P.I. and Nuzzo,M.I.
  Synthetic gene encoding human epidermal growth factor 2/neu antigen
  and uses thereof
JOURNAL    Patent: WO 2005012527-A 10 10-FEB-2005;
            Istituto di Ricerche di Biologia Molecolare P. Angeletti S.P.A.
            (IT)

FEATURES
source     Location/Qualifiers
1 .2028
   /organism="synthetic construct"
   /mol_type="unassigned DNA"
   /db_xref="taxon:32630"
   /note="HER2ECDTwt"

ORIGIN
Alignment Scores:
Pred. No.: 2e-10      Length: 2028
Score: 128.00        Matches: 23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%    Mismatches: 1
Query Match: 96.97%      Indels: 0
DB: 6                Gaps: 0

US-09-632-036F-6 (1-24) x CS020247 (1-2028)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGCACAACCAAGAGGTGACCGCGAGGACGCCAGCCAGCGCTGCGAGAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 12
CS075249
LOCUS      2061 bp      DNA      linear      PAT 05-MAY-2005
DEFINITION Sequence 15 from Patent WO2005033133.
ACCESSION CS075249
VERSION    CS075249.1 GI:63091615
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

REFERENCE
1 Shemesh,R., Oren,A., Rotman,G., Sela-Tavor,O., Walach,S.,
  Sameach-Greenwald,S., Beiman,M., Bahel,D. and Savitsky,K.
  Polynucleotides encoding novel ErbB-2 polypeptides and kits and
  methods using same
JOURNAL    Patent: WO 2005033133-A 15 14-APR-2005;
            Compugen Ltd. (IL)

FEATURES
source     Location/Qualifiers
1 .2061
   /organism="Homo sapiens"
   /mol_type="unassigned DNA"
   /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.03e-10   Length: 2061
Score: 128.00        Matches: 23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%    Mismatches: 1
Query Match: 96.97%      Indels: 0
DB: 6                Gaps: 0

US-09-632-036F-6 (1-24) x CS075249 (1-2061)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACAACCAAGAGGTGACAGGAGGTGACAGGAGTGGAAACACAGCGGTGTGAGAAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 13
CS085268
LOCUS      2086 bp      DNA      linear      PAT 25-MAY-2005
DEFINITION Sequence 10 from Patent WO2005039618.
ACCESSION CS085268
VERSION    CS085268.1 GI:66711225
KEYWORDS   unidentified
SOURCE     unidentified
ORGANISM   unclassified.

REFERENCE
1 Amici,A., Cavallo,F., Forni,G. and Marchini,C.
  Pl85 sp neu /sp -encoding DNA and therapeutical uses thereof
  Patent: WO 2005039618-A 10 06-MAY-2005;
  Indena S.p.A. (IT)

FEATURES
source     Location/Qualifiers
1
```

```
source
1. .2086
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 2.06e-10 Length: 2086
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CS085268 (1-2086)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 977 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1036

QY 21 LysProCysAla 24
Db 1037 AAGCCCTGTGCC 1048

RESULT 14
LOCUS CS085269 2086 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 11 from Patent WO2005039618.
ACCESSION CS085269
VERSION CS085269.1 GI:66711226
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
1
AUTHORS Amici,A., Cavallo,F., Forni,G. and Marchini,C.
TITLE P185 sp neu /sp -encoding DNA and therapeutical uses thereof
JOURNAL Indena S.p.A. (IT)
Patent: WO 2005039618-A 11 06-MAY-2005;
Indena S.p.A. (IT)

FEATURES
source
1. .2086
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 2.06e-10 Length: 2086
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CS085271 (1-2086)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 977 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1036

QY 21 LysProCysAla 24
Db 1037 AAGCCCTGTGCC 1048

RESULT 16
LOCUS CS085271 2086 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 13 from Patent WO2005039618.
ACCESSION CS085271
VERSION CS085271.1 GI:66711228
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
1
AUTHORS Amici,A., Cavallo,F., Forni,G. and Marchini,C.
TITLE P185 sp neu /sp -encoding DNA and therapeutical uses thereof
JOURNAL Indena S.p.A. (IT)
Patent: WO 2005039618-A 13 06-MAY-2005;
Indena S.p.A. (IT)

FEATURES
source
1. .2086
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 2.06e-10 Length: 2086
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CS085270 (1-2086)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 977 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1036

QY 21 LysProCysAla 24
Db 1037 AAGCCCTGTGCC 1048

RESULT 15
LOCUS CS085270 2086 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 12 from Patent WO2005039618.
ACCESSION CS085270
VERSION CS085270.1 GI:66711227
KEYWORDS
SOURCE unidentified
ORGANISM unclassified
```

LOCUS CS075235 2320 bp DNA linear PAT 05-MAY-2005  
DEFINITION Sequence 1 from Patent WO200503133.  
ACCESSION CS075235  
VERSION CS075235.1 GI:63091608  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Shemesh, R., Oren, A., Rotman, G., Sela-Tavor, O., Walach, S.,  
Sameah-Greenwald, S., Beiman, M., Eshel, D. and Savitsky, K.  
TITLE Polynucleotides encoding novel ErbB-2 polypeptides and kits and  
methods using same  
JOURNAL Patent: WO 200503133-A 1 14-APR-2005;  
Compugen Ltd. (IL)  
FEATURES  
source Location/Qualifiers  
1..2320  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.3e-10 Length: 2320  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x CS075235 (1-2320)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155  
QY 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167  
RESULT 18  
LOCUS AR082744 2385 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 3 from patent US 5976546.  
ACCESSION AR082744  
VERSION AR082744.1 GI:10009534  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2385)  
AUTHORS Laus, R., Ruegg, C.Landon. and Wu, H.  
TITLE Immunostimulatory compositions  
JOURNAL Patent: US 5976546-A 3 02-NOV-1999;  
FEATURES  
source Location/Qualifiers  
1..2385  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.37e-10 Length: 2385  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x AR082744 (1-2385)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20

Db 956 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1015  
QY 21 LysProCysAla 24  
Db 1016 AAGCCCTGTGCC 1027  
RESULT 19  
LOCUS AR099963 2385 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 3 from patent US 6080409.  
ACCESSION AR099963  
VERSION AR099963.1 GI:12810411  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2385)  
AUTHORS Laus, R., Ruegg, C.Landon. and Wu, H.  
TITLE Immunostimulatory method  
JOURNAL Patent: US 6080409-A 3 27-JUN-2000;  
FEATURES  
source Location/Qualifiers  
1..2385  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.37e-10 Length: 2385  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x AR099963 (1-2385)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 956 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1015  
QY 21 LysProCysAla 24  
Db 1016 AAGCCCTGTGCC 1027  
RESULT 20  
LOCUS AR143949 2385 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 3 from patent US 6210662.  
ACCESSION AR143949  
VERSION AR143949.1 GI:15105816  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2385)  
AUTHORS Laus, R., Ruegg, C.Landon. and Wu, H.  
TITLE Immunostimulatory composition  
JOURNAL Patent: US 6210662-A 3 03-APR-2001;  
FEATURES  
source Location/Qualifiers  
1..2385  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.37e-10 Length: 2385  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0



STAYNQLSSLSSESAVYCARWYMGQTTVTSSGTPGLDTHHTSGKPLDGBYFTL  
QIRFERFERFELNEALELDAQAGKEPGSGGSGTQVCTGDMKLRVLPASPEHL  
DMRLHYQCOVQOQNLLEYLPTNASLSFLQIOEVGYVLIANQVROVQLORLRI  
VRGTLQFDNVALALVLDNGDPLNNTPTVTGASPGGLRELQRLSLEILKGVILQRP  
CLQCQDTLWKDIFHKNQALTLIDTNRACUPCSPMCKSGWSSSDCOSLTV  
TVGASCPNPGRYTFGASCVTACPNYLSVDGCTCLVPHNQEVTAEQTCCEK  
TDTFESMPNPGRYTFGASCVTACPNYLSVDGCTCLVPHNQEVTAEQTCCEK  
CSKPCRCYGLGMEHLEVRVAVTSANTQEFAGCKKIFGSLAFILPESPDGPASNTAP  
LWQLOQVFTLEETGYLYISAWPDLPSVFNQLOVIRGRILHNGAYSLTLQGLG  
ISWGLRSLRELGLGALIHNTLCLFVHTVPDQLPENPHQALLTANRDESCVGE  
GLACHOLCARGHCWGPCTQCNCVSOFLRGQECVEECVVLQGLPREYNARHCLPCHP  
EQCPQNGSVTCFGEADQCACAHYKOPFFCVARCPGKVPDLSTYMPKPPDEEGAC  
QPCPINCTHSCVDLDDKCPAEQASPLTSGHHHHH"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,89e-10 Length: 2871  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x AX023363 (1-2871)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1828 CCCCTGCACACCAAGAGGTGCACAGCAGAGGTGGAACACAGCGGTGTGAGAAGTGCAGC 1887  
Qy 21 LysProCysAla 24  
Db 1888 AAGCCCTGTGCC 1899

## RESULT 24

AX05114  
LOCUS AX05114 3678 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 7 from Patent WO0240059.  
ACCESSION AX05114  
VERSION AX05114.1 GI:23386421

## KEYWORDS

Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

## REFERENCE

1 Mincheff,M.S., Loukinov,D.I. and Zoubak,S.

Methods and compositions for inducing cell-mediated immune

responses

Patent: WO 0240059-A 7 23-MAY-2002;

American Foundation for Biological Research Inc. (US) ; Mincheff,

Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)

Location/Qualifiers

1. .3678

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

7. .3678

/notes="unnamed protein product"

/codon\_start=1

/protein\_id="CAD48804.1"

/db\_xref="GI:23386422"

IQBQVGVLIHNVQVQVQLRIRVGTQLFDENYALVLDNGDPLNTPVTGASP  
GURELQRLSLEILKGVLIQRNPQLCYQDTILWKDIFHKNQALTLIDTNRSRAC  
HPCSPMCKSGRSGWSSSDCOSLTVTCAGCARCKGRLPTDCCHQACAGCTGPKHS  
DCLACLHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPNYLSVDV  
GSTCLVCLHNVQVTAEDGTQCEKSCPCARVCYGLGMEHLEVRVAVTSANTQEFAG  
CKKIFGSLAFILPESPDGPASNTAPLQEQLOVPELLEETGYLYISAWPDLPSLV  
FQNQLVIRGRIHNGAYSLTLQGLGISWLGRLSELGSLALIHNTLCLFVHTVPW  
DQLFRNPHQALLHNGAYSLTLQGLGISWLGRLSELGSLALIHNTLCLFVHTVPW  
VEECVRLQGLPREYNARHCLPCHPQCPQNGSVTCFGEADQCACAHYKOPFFCVA  
RCPGSGVKPDLSTYMPKPPDEEGACQPCPINCTHSCVDLDDKCPAEQASPLTSG

AVGILLVWVGVVGVGILIKRROOKIRKYTMRRLLQETELVEPLTPSGAMPNOQMIRI  
LKATELRKVKVLGSGAGTGYVKGWI PDGENVKIPVAIKVLRENTSPKANKEILDEAY  
VMKAGVSGVSRLLIGLICLTSTVQLVQIMPGYGLLDHVRENRLSGDILLWCMQOIA  
KGMWISDEVRLVHRDLAARNVLKSPNHVKITDFGLARLLDIDTETEHADGKVPKIKW  
MALESILURRFTHQSDVSYVTWELMTFGAKPYDGIIPAREIPDLLEKGBRLPQPII  
CTIDVYIMVMKWMIDSECRPRFRRELVSFSESRMARDQRFVVIQNEIDIGPASPLDSTF  
VRSILLEDDDLVDABEYLVPQGFCDPDPAGAGGVHHRHSSSTRSGGDLTLGL  
LEPSEBAPRSLPASPAGSDVDFDGLMGAAKGLSLPHTDPSLQRYSEDPTVPL  
PSETDGVVAPLTCSPOPEYVNDVRPQPSPREGPIPAARPGATLERAKTLSPGKN  
GVYKDVPAFGAVENPEYLTPOGGAAIQPHPPPAFSAFDNLYWDQDPBERGAPPST  
FKGTPPTAENPEYLGLDVFPV"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,75e-10 Length: 3678  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x AX05114 (1-3678)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 856 CCCCTGCACACCAAGAGGTGCACAGCAGAGGTGGAACACAGCGGTGTGAGAAGTGCAGC 915  
Qy 21 LysProCysAla 24  
Db 916 AAGCCCTGTGCC 927

## RESULT 25

AR034479  
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869445.  
ACCESSION AR034479  
VERSION AR034479.1 GI:5950084

## KEYWORDS

Unknown.

## SOURCE

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 3768)

AUTHORS Cheever,M.A. and Disis,M.L.

TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein

JOURNAL

Patent: US 5869445-A 1 09-FEB-1999;

FEATURES

Location/Qualifiers

1. .3768

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x AR034479 (1-3768)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACACCAAGAGGTGCACAGCAGAGGTGGAACACAGCGGTGTGAGAAGTGCAGC 1005  
Qy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

## RESULT 26

BD224136  
LOCUS BD224136 3768 bp DNA linear PAT 17-JUL-2003  
DEFINITION Novel method for vaccine injection for therapeutic purpose.

```
ACCESSION BD224136
VERSION BD224136.1 GI:33033906
KEYWORDS JP 2002526419-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3768)
AUTHORS Steinha.L., Mouritsen,S., Nielsen,K.G., Haaning,J., Leach,D.,
DALUM,I., Gautam,A., Birk,P. and Karlsson,G.
TITLE Novel method for vaccine injection for therapeutic purpose
JOURNAL Patent: JP 2002526419-A 2 20-AUG-2002;
M AND E BIOTECH AS
COMMENT OS Homo sapiens (human)
PN JP 2002526419-A/2
PD 20-AUG-2002
PF 05-OCT-1999 JP 2000573386
PR 05-OCT-1998 DK PA 199801261,20-OCT-1998 US 60/105011 PI
LUCILLA STEINAA,SOREN MOURITSEN,KLAUS GREGORIUS NIELSEN,JESPER
PI HAANING,
PC DANA LEACH, IBEN DALUM,ANAND GAUTAM,PETER BIRK GUNILLA KARLSSON
PC A61K39/00,A61K39/39,A61P15/00,A61P35/00,C07K14/47,C07K16/18//
PC C12N15/09,
PC C12N15/00
CC Novel method for vaccine injection for therapeutic purpose FH
KEY Location/Qualifiers
FT CDS (1)..(3768).
Alignment Scores:
Pred. No.: 3,85e-10 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0
US-09-632-036F-6 (1-24) x BD224136 (1-3768)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACAGAGGTGACAGCAGAGGTGGAACACACAGCGGTGAGAAGTGCAGC 1005
QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017
RESULT 27
BD267514 3768 bp DNA linear PAT 17-JUL-2003
LOCUS HER-2/neu fusion proteins.
DEFINITION BD267514
ACCESSION BD267514
VERSION BD267514.1 GI:33077282
KEYWORDS JP 2002535004-A/1.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3768)
AUTHORS Cheever,M.A. and Gheysen,D.
TITLE HER-2/neu fusion proteins
JOURNAL Patent: JP 2002535004-A 1 22-OCT-2002;
CORIXA CORP. SMITHKLINE BEECHAM PLC
COMMENT OS Homo sapiens (human)
PN JP 2002535004-A/1
PD 22-OCT-2002
PF 28-JAN-2000 JP 2000596141
PR 29-JAN-1999 US 60/117976
PI MARTIN A CHEEVER, DIRK GHEYSEN
PC C12N15/09,A61K35/76,A61K38/00,A61K39/00,A61K48/00,A61P35/00,
PC A61P37/04,
PC C07K14/82,C07K19/00,C12N5/10,C12P21/02//(C12P21/02,C12R1:91),
PC C12N15/00,
PC C12N5/00,A61K37/02
CC human HER-2/neu protein
CC extracellular domain (ECD) of human HER-2/neu CC
intracellular domain (ICD) of human HER-2/neu CC phosphorylation
domain (PD) of human HER-2/neu CC preferred portion of the
phosphorylation domain (delta PD) of CC
CC HER-2/neu
FH Key Location/Qualifiers
FT CDS (1)..(3768)
FT misc feature (1)..(1959)
FT misc feature (2026)..(3765)
FT misc feature (2968)..(3765)
FT misc feature (2968)..(3144).
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 3,85e-10 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0
US-09-632-036F-6 (1-24) x BD267514 (1-3768)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACAGAGGTGACAGCAGAGGTGGAACACACAGCGGTGAGAAGTGCAGC 1005
QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017
RESULT 28
CQ815384 3768 bp DNA linear PAT 02-JUN-2004
LOCUS Sequence 1 from Patent EP1418235.
DEFINITION CQ815384
ACCESSION CQ815384
VERSION CQ815384.1 GI:47971074
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Cheever,M. and Disis,M.L.
TITLE Intracellular domain of the her-2/neu protein for prevention or
treatment of malignancies
JOURNAL Patent: EP 1418235-A 1 12-MAY-2004;
University of Washington (US)
FEATURES
source
CDS
1..3768
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAG32807.1"
/db_xref="GI:47971075"
/translation="MELAALCRWGLLLALLPFGAASVQVCTGTDKMLRLPASPETHLD
MLRHLYQCGVQVQGNLELTPLPTNASLSFLQDIQEVQGVLYIAHNQVRQVPLQRLRIV
```



RGTQLPEDNYALAVLDNGDPLNNNTTPTVTGASPGGLRELOLRSLTEILKGGVLQIRNPQ  
LCYQDTILWKDI1PHKNQALHTLIDTNRSRACHPCSPMCKGSRGWSSDSCOSLRPT  
VCAGCARCKPLPTDCTHQCAAGCTGPKHSDCLACHFNHSGICELHCPALVTYNT  
KTESPMNPSGRVYTFEGASCVTACPNVYLTSDVGSCTLVCPLNHOEVTAEQGTQCEKC  
SKPCARVYGLGHEHLREVAVTSANIOEFAGCKKIFGSLAFIPESFGDPSANTAPL  
QPOLQVFETLBEITGYLYISAMPDSDUDSVFQNLQVIRGIRLHNGAYSLTLQGLGI  
SWLGRSLRELGLSLAIHNTLHLCFVHTVPMQLFERNPHOALLHTANREDECVSEB  
LACHQLCARGHGPQGTQCVNCSQFLRGQECVEECRLQGLPREYVNAHCLPCHPE  
CPQNGSVTCFGEADOCVACAHYKDPFCPCPGKVPDLSYMPYTWKPFDEBGAQ  
PCPINTCHSCVDLDDKCCPAORASPLTSLISAVVGLLVVVLGVVFGIILKRQOKI  
RKYTMRELLAETELVELPTSGAMPNQAOMRILKETBLRKVKVLGSGAFGTVYKGIWI  
PDENVKIPVAIKVLRNTSPKANKEILDSEAYMAGVGSYPVSRLLGICUTSTVLVT  
QLMPYGLLDHVRNRLGSLQDLNMCQIMAGMSYLEDVRLVHRDLAARNVLKSP  
LHMTKIDFLGLARLLIDIDETBYHADGGKVP1KMMALESILRRRFTHOSDVMYSYVWME  
LMTFGAKYDPIPAEPTDLEKGERLPQPICTIDVYIMVWKMMIDSDCRFRREL  
VSEFSRMDPQRFVTONEDLGPASPLDSTFVRSLEDDDDMGDLVDAEYLYVQOQF  
FDPDAPAGMWHHRSSSTSGSGDLTLGLEPSEEAAPRPLAPSEBAGSDVDFG  
DLGMAKAGLQSLTHDPSPLQRYSEDTVPLPSETDGYVAPLTSPQBPVYNQPDVR  
POPPSPREGELPAAPAGATLERPKTLSPGKNVVKDVFAGFAGAVENPEYLTPOGGAA  
POPHPPAFSPFDNLYYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVFPV"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CQ815384 (1-3768)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTCGACAAACCAAGAGGTGACAGCAGAGAGTGAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

## RESULT 29

CQ841302  
LOCUS CQ841302 3768 bp DNA linear PAT 02-AUG-2004  
DEFINITION Sequence 1 from Patent WO2004061105.  
ACCESSION CQ841302  
VERSION CQ841302.1 GI:50893095  
KEYWORDS Macaca mulatta (rhesus monkey)  
SOURCE Macaca mulatta  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopitheciidae; Cercopitheciinae; Macaca.

REFERENCE 1  
AUTHORS Ciliberto, G.I., Lahm, A.I., la Monica, N.I., Monaci, P.I. and Nuzzo, M.I.  
TITLE Rhesus her2/neu, nucleotides encoding same, and uses thereof  
JOURNAL Patent: WO 2004061105-A 1 22-JUL-2004;  
INSTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.  
(IT)

## FEATURES

source  
1. .3768 Location/Qualifiers  
/organism="Macaca mulatta"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9544"  
misc\_feature 1. .3768  
/notes="R = A or G"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1

Query Match:  
DB: 96.97% Indels: 0  
Gaps: 0

US-09-632-036F-6 (1-24) x CQ841302 (1-3768)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTCGACAAACCAAGAGGTGACAGCAGAGTGAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

## RESULT 30

CQ841341  
LOCUS CQ841341 3768 bp DNA linear PAT 02-AUG-2004  
DEFINITION Sequence 40 from Patent WO2004061105.  
ACCESSION CQ841341  
VERSION CQ841341.1 GI:50893133  
KEYWORDS Macaca mulatta (rhesus monkey)  
SOURCE Macaca mulatta  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopitheciidae; Cercopitheciinae; Macaca.

REFERENCE 1

AUTHORS Ciliberto, G.I., Lahm, A.I., la Monica, N.I., Monaci, P.I. and Nuzzo, M.I.  
TITLE Rhesus her2/neu, nucleotides encoding same, and uses thereof  
JOURNAL Patent: WO 2004061105-A 40 22-JUL-2004;  
INSTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.  
(IT)

## FEATURES

source  
1. .3768 Location/Qualifiers  
/organism="Macaca mulatta"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9544"  
misc\_feature 1. .3768  
/notes="R = A or G  
Y = C or T"

## ORIGIN

Alignment Scores:  
Pred. No.: 3,85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CQ841341 (1-3768)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTCGACAAACCAAGAGGTGACAGCAGAGTGAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

## RESULT 31

CQ841343  
LOCUS CQ841343 3768 bp DNA linear PAT 02-AUG-2004  
DEFINITION Sequence 42 from Patent WO2004061105.  
ACCESSION CQ841343  
VERSION CQ841343.1 GI:50893134  
KEYWORDS Macaca mulatta (rhesus monkey)  
SOURCE Macaca mulatta  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopitheciidae; Cercopitheciinae; Macaca.

REFERENCE 1

AUTHORS Ciliberto,G.I., Lahm,A.I., la Monica,N.I., Monaci,P.I. and Nuzzo,M.I.  
TITLE Rhesus her2/neu, nucleotides encoding same, and uses thereof  
JOURNAL Patent: WO 2004061105-A 42 22-JUL-2004;  
ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.  
(IT)

FEATURES  
source Location/Qualifiers  
1..3768  
/organism="Macaca mulatta"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9544"

ORIGIN

Alignment Scores:

Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CQ841343 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTGTCACAAACGAGGTGACAGCGGAGGACGAACACAGCGATGTGAGAAGTGCAGC 1005  
QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

RESULT 32  
CQ8411344  
LOCUS CQ8411344 3768 bp DNA linear PAT 02-AUG-2004  
DEFINITION Sequence 43 from Patent WO2004061105.  
ACCESSION CQ8411344  
VERSION CQ8411344.1 GI:50893135  
KEYWORDS  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE 1  
AUTHORS Ciliberto,G.I., Lahm,A.I., la Monica,N.I., Monaci,P.I. and Nuzzo,M.I.  
TITLE Rhesus her2/neu, nucleotides encoding same, and uses thereof  
JOURNAL Patent: WO 2004061105-A 43 22-JUL-2004;  
ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.  
(IT)

FEATURES  
source Location/Qualifiers  
1..3768  
/organism="Macaca mulatta"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9544"

ORIGIN

Alignment Scores:

Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CQ841344 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTGTCACAAACGAGGTGACAGCGGAGGACGAACACAGCGATGTGAGAAGTGCAGC 1005  
QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

Db 1006 AAGCCCTGTGCC 1017  
RESULT 33  
CQ007994  
LOCUS CQ007994 3768 bp DNA linear PAT 11-FEB-2005  
DEFINITION Sequence 3 from Patent EP1502602.  
ACCESSION CQ007994  
VERSION CQ007994.1 GI:59668913  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
Steinaa,L., Mouritsen,S., Nielsen,K.G., Haaning,J., Leach,D.,  
Dalum,I., Gautam,A., Rasmussen,P.B. and Karlsson,G.  
Methods for therapeutic vaccination  
Patent: EP 1502602-A 3 02-FEB-2005;  
Pharmexa A/S (DK)  
Location/Qualifiers  
1..3768  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..3768  
/note="unnamed protein product"

CDS

/protein\_id="CAI53479.1"  
/translation="MELAALCRWGLLLALLPPGAASTQVCTGTMKLRLPASPETHLD  
MLRHLYQGVVQVQNGLELYPLNTNASLFLQDIQEVGVYLIHAHQVQVPLQRLRIV  
RGTOIFEDNYALAVLDNGDPLNNTPTVTGASPGRLQRLSLTLILKGVLIQRNPQ  
LCYODTILWKDIFHKNNOLALTLDITNSRACHPCS PMCKSGKSGCESSEDQSLTRT  
VCAGSACRKGPLPTDCCEHQCAAGCTGPKHSDCLACILHFNHSGICELHCPALVTYT  
DTFESMPNPEGRITFTGASCVTACTPNTSLTDVSGCTLVCPLNQVETADGTCRCEKC  
SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKKI FGS LAF L P S F D G D P A S N T A P L  
QPEQLQVFTLEETITGYLISAWPDSLPLDSVFNQNLVIRGRILHNHGYSLTLQGLGI  
SWLGLRSLRELGSLALIHNTHLCFVHTVPMWDLFRNPHQALLTANRPEDECVGEG  
LACHOLCARGHCWGPGTQVCNCSQFLGQECVEBCRVLQGLPREVYNARHCLPCHPE  
CQPNQSVTCFGEADQCVACAHYKDPFCVACRCPGVKPDLSYMPIKWFPEEGACQ  
PCYINTHSCVDLDDKCPAQRASPLISVAVVGILLVVGIVVFGVGLIKRQOKI  
PKYTWRLQLTELVEPLTPSGAPNQAMRILKETLRKKVKVGLSGAGFTVYKGIWI  
OLMPYGCLLDHVRENRLGSDLLNWCQIAKMSYLEDLVRLHVRDLAARNLVKSP  
NHVKITDFGLARLLDIDETEVHADGKVPKIKWMALESILRRRFTHQSDVMSYGVTVWE  
LMTFGAKPYDGIIPAREIPDLLEKGRLPQPPICTIDVYIMVUKWMIQSECRPREL  
VSEFRMARDPQRFVVIQNEGLPASPDLSTFYRSLLEDDDDMDGLVDAREYLVPOQGF  
FCPDPAAGAGVMVHRHRSSTSRSGDGLTLGLEPSEEEAPRSLPASEGASDVDFG  
DLGMAKGLQSLPLTHDPSPQRYSGEDFTVPLPSETDGYAELTCSPOEYVNPQDVR  
PQPPSPREGPLPAARPAAGTAKTLSPGKNGVVKDVFAGCGAVENPEYLTPOGGAA  
PQPHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPPTAENPEYLGLDVVPV"

ORIGIN

Alignment Scores:  
Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CQ007994 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCTGTCACAAACGAGGTGACAGAGGTGACAGCGGTGTGAGAAGTGCAGC 1005  
QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 34

CS020239  
LOCUS CS020239 3768 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 1 from Patent WO2005012527.  
ACCESSION CS020239  
VERSION CS020239.1 GI:60220901  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Gallo,P.I., Monaci,P.I. and Nuzzo,M.I.  
TITLE Synthetic gene encoding human epidermal growth factor 2/neu antigen and uses thereof  
JOURNAL Patent: WO 2005012527-A 1 10-FEB-2005;  
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.P.A.  
(IT)  
FEATURES  
source Location/Qualifiers  
1..3768  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="HER2opt"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x CS020239 (1-3768)  
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACAACACGAGAGGTGACCGCGGACGCGGACGCCGCCCGCTCGGAGAGTGCAGC 1005  
Qy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 35  
LOCUS CS104222 3768 bp DNA linear PAT 10-JUN-2005  
DEFINITION Sequence 2 from Patent WO2005049868.  
ACCESSION CS104222  
VERSION CS104222.1 GI:67512453  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Mehta,A. and Trotta,C.R.  
TITLE Methods and agents for screening for compounds capable of modulating her2 expression  
JOURNAL Patent: WO 2005049868-A 2 02-JUN-2005;  
PCT Therapeutics Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..3768  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0

DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x CS104222 (1-3768)  
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACAACACGAGAGGTGACGAGAGGTGGAACACACAGCGGTGTGAGAGTGCAGC 1005  
Qy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 36  
LOCUS AR409603 3768 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 2 from patent US 6632979.  
ACCESSION AR409603  
VERSION AR409603.1 GI:40160579  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3768)  
AUTHORS Erickson,S., King,K. and Schwall,R.  
TITLE Rodent HER2 tumor model  
JOURNAL Patent: US 6632979-A 2 14-OCT-2003;  
Genentech, Inc.; South San Francisco, CA  
FEATURES  
source Location/Qualifiers  
1..3768  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.85e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x AR409603 (1-3768)  
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACAACACGAGAGGTGACGAGAGGTGGAACACACAGCGGTGTGAGAGTGCAGC 1005  
Qy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 37  
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 2 from Patent WO0100244.  
ACCESSION AX060704  
VERSION AX060704.1 GI:12406101  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Erickson,S. and Schwall,R.  
TITLE Methods of treatment using anti-erbB antibody-maytansinoid conjugates  
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;  
Genentech, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..3768  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"

ORIGIN		/db_xref="taxon:9606"	
Alignment Scores:			
Pred. No.:	3.85e-10	Length:	3768
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	6	Gaps:	0
US-09-632-036F-6 (1-24) x AX060704 (1-3768)			
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20			
Db	946	CCCCTGCAACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC	1005
QY 21 LysProCysAla 24			
Db	1006	AAGCCCTGTGCC	1017
RESULT 38			
AX201817	AX201817	3768 bp	DNA linear PAT 30-AUG-2001
LOCUS	Sequence 1 from Patent WO0153463.		
DEFINITION	AX201817		
ACCESSION	AX201817		
VERSION	AX201817.1	GI:15391666	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Cheever, M.A. and Hand-Zimmermann, S.		
TITLE	Compounds and methods for prevention and treatment of her-2/ neu associated malignancies		
JOURNAL	Patent: WO 0153463-A 1 26-JUL-2001;		
	CORIXA CORPORATION (US)		
FEATURES	Location/Qualifiers		
source	1..3768		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	1..3768		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC60126.1"		
	/db_xref="GI:15391667"		
	/translation="MELAALCRWGLLLALLPGGAASVQVCTGTGDMKLRLPASPE		
	THLD		
	MRHLYGQGVVQGNLELTYLPTNASLFLQDIEVQGVYLIAHNOVROVPLQRLRIV		
	RGTQLFEDNYALVLDNGDPLNNTPTVTGASPGGLRELQLRSITELTKGVLIQRNPQ		
	LCYQDTILWKDIFPHKNQALTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRT		
	VCAGGCAKRGPLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT		
	DTFESMPNPEGRTYFGASCVTACPYNYLSDVGSCTLVCPLHQEVTAEDGTQRCCK		
	SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPL		
	QPQLQVFTLEETICGLYISAMPDLSPLSVFQNLQVIRGRILHNGAYSLTLQGLGI		
	SMGLRSRLRELGSGLALIHNNTHLCFVHTVPMDLFRNPHOALLTANRPECVCEG		
	LACHOLCARGHCWGCPQTCVNCQFRLGQECVBEICVLQGLPREYVNAHCLPCHPE		
	CQPNQSVTCFGEADQCVACAHYKPPPCVACRCPGVPDLSYMEIWFKPFBEAGCQ		
	PCPINTHSCVDLDDKGFPAEQRASPLTSIIISAVVGILLVVLVGVVGLIKRRQKI		
	RKYMTRMLQIETLEVEPLTPSGAMPNQAMRILKETELRKVKVLSGAFGTYYKGIWI		
	PLQENVKI PVAI KVLRENTSPKAKEIIDEAYVMAGVGSPPYVSLIGI CLTSTVQLVT		
	OLMPYGCLLDHRVNRGRGLSQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLXSP		
	NHWKTFDGLARLLDIDETEYHADGCKVPIKWMALSIILRRSFTHQSDVNSYGVTVWE		
	LMFTGAKPYDGPAREIPDLLEKGERLPQPPCTIDVYIMVKMWIDSECRPRFREL		
	VSEFSKMAKDQRFVVIQNEIDGSPASPLDSTFYRSLLEDDMDGLVDASEYLVPOQGF		
	FCDDPAPAGGMVHRRHSSSTRSGGDLTLGLEPSEEAAPRSPAPSGAGSDVPDG		
	DLGMAAGLQSLPHTDPSPLORYSEDPTVPLPSETDGYVAPLTCSPOQEVYNQDVR		
	POPPSPREGPLPAARPAAGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGA		
	PQHPHPPAPSPAFDNLVYWDQDPPERGAFFTFTKGTPTAENPEYLGLDVVPV"		
misc_feature			
	1..1959		
	/note="extracellular domain (ECD) of human Her-2/neu"		
misc_feature			
	2026..3765		
	/note="intracellular domain (ICD) of human Her-2/neu"		
misc_feature			
	2968..3765		
	/note="phosphorylation domain (PD) of human Her-2/neu"		

ORIGIN		/db_xref="taxon:9606"	
Alignment Scores:			
Pred. No.:	3.85e-10	Length:	3768
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	6	Gaps:	0
US-09-632-036F-6 (1-24) x AX201817 (1-3768)			
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20			
Db	946	CCCCTGCAACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC	1005
QY 21 LysProCysAla 24			
Db	1006	AAGCCCTGTGCC	1017
RESULT 39			
AX380923	AX380923	3768 bp	DNA linear PAT 18-MAR-2002
LOCUS	Sequence 9 from Patent WO0212341.		
DEFINITION	AX380923		
ACCESSION	AX380923		
VERSION	AX380923.1	GI:19575767	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Cheever, M.A. and Gheyssen, D.		
TITLE	Her-2/neu fusion proteins		
JOURNAL	Patent: WO 0212341-A 9 14-FEB-2002;		
	CORIXA CORPORATION (US); SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)		
FEATURES	Location/Qualifiers		
source	1..3768		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	1..3768		
	/note="unnamed protein product; human Her-2/neu protein"		
	/codon_start=1		
	/protein_id="CAD28392.1"		
	/db_xref="GI:19575768"		
	/translation="MELAALCRWGLLLALLPGGAASVQVCTGTGDMKLRLPASPE		
	THLD		
	MRHLYGQGVVQGNLELTYLPTNASLFLQDIEVQGVYLIAHNOVROVPLQRLRIV		
	RGTQLFEDNYALVLDNGDPLNNTPTVTGASPGGLRELQLRSITELTKGVLIQRNPQ		
	LCYQDTILWKDIFPHKNQALTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRT		
	VCAGGCAKRGPLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT		
	DTFESMPNPEGRTYFGASCVTACPYNYLSDVGSCTLVCPLHQEVTAEDGTQRCCK		
	SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPL		
	QPQLQVFTLEETICGLYISAMPDLSPLSVFQNLQVIRGRILHNGAYSLTLQGLGI		
	SMGLRSRLRELGSGLALIHNNTHLCFVHTVPMDLFRNPHOALLTANRPECVCEG		
	LACHOLCARGHCWGCPQTCVNCQFRLGQECVBEICVLQGLPREYVNAHCLPCHPE		
	CQPNQSVTCFGEADQCVACAHYKPPPCVACRCPGVPDLSYMEIWFKPFBEAGCQ		
	PCPINTHSCVDLDDKGFPAEQRASPLTSIIISAVVGILLVVLVGVVGLIKRRQKI		
	RKYMTRMLQIETLEVEPLTPSGAMPNQAMRILKETELRKVKVLSGAFGTYYKGIWI		
	PLQENVKI PVAI KVLRENTSPKAKEIIDEAYVMAGVGSPPYVSLIGI CLTSTVQLVT		
	OLMPYGCLLDHRVNRGRGLSQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLXSP		
	NHWKTFDGLARLLDIDETEYHADGCKVPIKWMALSIILRRSFTHQSDVNSYGVTVWE		
	LMFTGAKPYDGPAREIPDLLEKGERLPQPPCTIDVYIMVKMWIDSECRPRFREL		
	VSEFSKMAKDQRFVVIQNEIDGSPASPLDSTFYRSLLEDDMDGLVDASEYLVPOQGF		
	FCDDPAPAGGMVHRRHSSSTRSGGDLTLGLEPSEEAAPRSPAPSGAGSDVPDG		
	DLGMAAGLQSLPHTDPSPLORYSEDPTVPLPSETDGYVAPLTCSPOQEVYNQDVR		
	POPPSPREGPLPAARPAAGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGA		
	PQHPHPPAPSPAFDNLVYWDQDPPERGAFFTFTKGTPTAENPEYLGLDVVPV"		
misc_feature			
	1..1959		
	/note="extracellular domain (ECD) of human Her-2/neu"		
misc_feature			
	2026..3765		
	/note="intracellular domain (ICD) of human Her-2/neu"		
misc_feature			
	2968..3765		
	/note="phosphorylation domain (PD) of human Her-2/neu"		

```
misc_feature 2968..3144
/note="preferred portion of the phosphorylation domain
(delta PD) of human Her-2/neu"

ORIGIN
Alignment Scores:
Pred. No.: 3,85e-10 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x AX380923 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACCAAGAGGTGACAGAGGATGGACACAGCGGTGTGAGAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 40
AX384604
LOCUS AX384604 3768 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0214503.
ACCESSION AX384604
VERSION AX384604.1 GI:19577806
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
Patent: WO 0214503-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD28577.1"
/db_xref="GI:19577807"
/translation="MELALACRWGLLLALLPGCAASTQVCTGDMKRLRPASPETHLD
MLRHLVGGCVQGNLELYLPTNASLFLQDIOEVQGYVLIHAHQVQVPLQRLRV
RGTFEDFNQYALVLDNGDPLNTTDTNRSRACHPCSPCKGSRWGESSEDQSLTRT
LCQDTILMKDILPHKNNQALITIDTNRACHPCSPCKGSRWGESSEDQSLTRT
VCAGGARCKGKPLPTDCCHEOCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNT
DTFESMNPPEGRTFCGSCVTACPNYVLTSDVGSCTLVCPLNQEVTAEDGTQCEKC
SKFCARVCLGMEHLREVRNAVTSANIQFAGCKIFGSLAFLPESFDGDPASNTAPL
QPQLQVFTEITGLYIISAWPDSLDPSVFNQVIRGRILHNGAYS LTLQGLGI
SWLGRSLRALIHHNTHLCFVHTVPMDLPNPHQALLHTANRPEDECVGEG
LACHQLCARGHCGPGTCVNCVSQFLRGQECVEECRVLQGLPREVYNARHCLPCHPE
COPQNSGVTFCPEADQCVACAHVKDPFECVARCPGKVPDL SYMPIWKEPDEEGACQ
PCPINTHSCVDLDDGCPAEQASPLTSIIISAVGILVVLGVVFGIILKRRQKI
RKYTRMLLQETELVELPTSPGAMPNQAMRILKETELRKVKVLGSGAFGTVYKGIWI
PGENYKI PVAIKRENTSPKANKIIDEAYVMAGVSPYVSRLLGICLTSTVQLVT-
QLMPYGLLDHVRNENRGLSQDLNWCQIAKMSYLEDRVLRDLAARNVVKSP
NHVKITDFGLRLDIDETEHADGKVPILKMALESILRRRFTHSDVMSYGVTVME
LMTFGAKPVDGIPAREIPDLLEKGERLPQPICTIDVYIMVWKMWIDSECRPREL
VSEFSRMADPQRPVVIQNEIDLGPASPLSTFYRSLDEDDMDGLVDABEYLVPOQGF
FCPDAPAGAGMHHRRHSSSTRSGGDLTLGLEPSEEA PRSPLAPSEGAGSDVPDG
DLGMAAGKQLPHTDPSLPQRYSEDPTVLPFSETDGYVAPLTCSPQPEYVNPQDVR
POPPSPREGPLPAARPAAGATLERPKTLPSPKNGVVKDVFAGGAVENPEYLTPOGGA
```

```
ORIGIN
Alignment Scores:
Pred. No.: 3,85e-10 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x AX384604 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACCAAGAGGTGACAGAGGATGGACACAGCGGTGTGAGAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

Search completed: December 11, 2005, 00:00:07
Job time : 2213.91 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 21:10:38 ; Search time 283.091 Seconds  
(without alignments)

565.022 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLHNQEVTAEDGTQRAEKCSPCA 24

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFio\_spool/US09632036/runat\_02122005\_104140\_25801/app\_query.fasta\_1.398  
-DB=N\_Geneseq -QMT=Fastcap -SUFFIX=rrng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09632036@cgn\_1\_1\_714 @runat\_02122005\_104140\_25801 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	97.0	471	6	ABK29542
2	128	97.0	670	10	ACD96468
3	128	97.0	1219	14	ADZ69875
4	128	97.0	1255	11	ADM29353

5	128	97.0	1260	10	ADF16383
6	128	97.0	1320	14	ADZ47818
7	128	97.0	1459	14	ADZ69874
8	128	97.0	1555	8	ACC83226
9	128	97.0	1699	14	ADZ69873
10	128	97.0	1872	2	AAQ06828
11	128	97.0	1939	14	ADZ69872
12	128	97.0	1944	14	ADZ47796
13	128	97.0	1956	12	ADJ57171
14	128	97.0	2028	14	ADM87396
15	128	97.0	2028	14	ADM87397
16	128	97.0	2052	12	ADJ57170
17	128	97.0	2061	14	ADZ47808
18	128	97.0	2086	14	ADZ69880
19	128	97.0	2086	14	ADZ69879
20	128	97.0	2086	14	ADZ69881
21	128	97.0	2086	14	ADZ69882
22	128	97.0	2132	10	ADG17937
23	128	97.0	2149	10	ADG17943
24	128	97.0	2164	10	ADG17939
25	128	97.0	2261	11	ADM29355
26	128	97.0	2320	14	ADZ47794
27	128	97.0	2385	2	AAT72725
28	128	97.0	2437	14	ADZ47806
29	128	97.0	2802	13	ADT50728
30	128	97.0	2871	3	AAZ50586
31	128	97.0	3188	11	ADM29351
32	128	97.0	3600	3	AAA89736
33	128	97.0	3678	6	ABK86207
34	128	97.0	3765	10	ADB67620
35	128	97.0	3765	14	ADZ72379
36	128	97.0	3768	2	AAT40739
37	128	97.0	3768	2	AAx01912
38	128	97.0	3768	3	AAA09455
39	128	97.0	3768	5	AAH23392
40	128	97.0	3768	6	ABV78168
41	128	97.0	3768	6	ABZ35744
42	128	97.0	3768	6	AA032743
43	128	97.0	3768	6	ABK10730
44	128	97.0	3768	6	AA043935
45	128	97.0	3768	6	ABX09987
46	128	97.0	3768	6	AA043986
47	128	97.0	3768	6	ABK14058
48	128	97.0	3768	6	ABA92250
49	128	97.0	3768	6	ABU91709
50	128	97.0	3768	8	ACC57649
51	128	97.0	3768	13	ADQ75970
52	128	97.0	3768	13	ADQ75929
53	128	97.0	3768	13	ADQ75968
54	128	97.0	3768	13	ADQ75971
55	128	97.0	3768	14	ADV43219
56	128	97.0	3768	14	ADM87389
57	128	97.0	3768	14	AEA47578
58	128	97.0	3778	14	ADM87398
59	128	97.0	4299	2	AAQ46083
60	128	97.0	4472	3	AAA14812
61	128	97.0	4473	2	ABQ76220
62	128	97.0	4473	2	AAZ31071
63	128	97.0	4473	6	ABZ34969
64	128	97.0	4473	6	ABZ38904
65	128	97.0	4473	8	ACC69999
66	128	97.0	4473	10	ADC35148
67	128	97.0	4473	10	ADD25483
68	128	97.0	4473	12	ADQ17192
69	128	97.0	4478	12	ADQ83974
70	128	97.0	4529	12	ADJ57169
71	128	97.0	4530	2	AAT01585
72	128	97.0	4530	2	AAT71253
73	128	97.0	4530	3	AAZ60815
74	128	97.0	4530	4	ABN19731
75	128	97.0	4530	6	ABN85585
76	128	97.0	4530	6	ABZ35012
77	128	97.0	4530	6	ABV94128

78 128 97.0 4530 6 ABK83918  
79 128 97.0 4530 8 ACC50139  
80 128 97.0 4530 8 ABQ83856  
81 128 97.0 4530 9 AAD58073  
82 128 97.0 4530 10 ADC09594  
83 128 97.0 4530 12 ADH13161  
84 128 97.0 4530 12 ADJ32564  
85 128 97.0 4530 12 ADM72832  
86 128 97.0 4530 12 ADO20008  
87 128 97.0 4530 12 ADG29633  
88 128 97.0 4530 13 ACN40176  
89 128 97.0 4530 13 ADR83426  
90 128 97.0 4530 13 ADW44364  
91 128 97.0 4530 14 ADW28639  
92 128 97.0 4530 14 ADY61191  
93 128 97.0 4530 14 ADZ09642  
94 128 97.0 4530 14 AEA15048  
95 128 97.0 4530 14 AEA08354  
96 128 97.0 4559 13 ADT50724  
97 128 97.0 4606 11 ACN89746  
98 128 97.0 4642 11 ACN89746  
99 128 97.0 4647 14 ADZ47802  
100 128 97.0 4806 14 ADY30408

## ALIGNMENTS

RESULT 1  
ABK29542  
ID ABK29542 standard; cDNA; 471 BP.  
XX  
AC ABK29542;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Colon adenocarcinoma-specific cDNA #68.  
XX  
KW Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200196389-A2.  
XX  
PD 20-DEC-2001.  
XX  
XX 07-JUN-2001; 2001WO-US018574.  
XX  
PR 09-JUN-2000; 2000US-0210667P.  
PR 22-NOV-2000; 2000US-0252614P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Meagher MJ, King GE, Xu J, Secrist H;  
XX  
XX WPI; 2002-098052/13.  
XX  
XX New isolated polynucleotide encoding a polypeptide comprising a portion  
PT of colon tumor protein, for detection, diagnosis and therapy of human  
PT colon cancer.  
XX  
PS Claim 1; Page 133; 211pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (I) encoding a  
CC polypeptide (II) comprising a portion of a colon tumour protein. A new  
CC oligonucleotide (III) that hybridises to (I) is useful for determining  
CC the presence of a cancer in a patient. (II) or antigen presenting cells  
CC expressing (I) are useful for stimulating and/or expanding T cells  
CC specific for a tumour protein, by contacting T cells with (I), (II) or  
CC antigen-presenting cells that express (I), (II), or antigen  
CC presenting cells that express (II) are useful for treating colon cancer  
CC in a patient by incubating CD4+ and/or CD8+ T cells isolated from a  
CC patient with (I), (II), or antigen presenting cells that express (II), so  
CC that T cells proliferate, and administering to the patient an effective

CC amount of the proliferated T cells, thus inhibiting the development of a  
CC cancer in the patient. A new composition is useful for stimulating an  
CC immune response in a patient. (I) or (II) is useful in vaccines and  
CC pharmaceutical compositions for prevention and treatment of colon cancer  
CC and for the diagnosis and monitoring of the cancers. (I), (II) or an  
CC antibody against (II) is useful for detection, diagnosis and/or therapy  
CC of human colon cancer. (I) is useful as a probe or primer for nucleic  
CC acid hybridisation, and in the design and preparation of ribozyme  
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-  
CC ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of  
CC the invention

SQ Sequence 471 BP; 104 A; 136 C; 135 G; 96 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,36e-11 Length: 471  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: Gaps: 0

US-09-632-036F-6 (1-24) x ABK29542 (1-471)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 65 CCCCCTGCACACCAAGAGGTGACAGCAGAGATGGAACACACGCGGTGTGAGAAGTGCAGC 124  
|||  
QY 21 LysProCysAla 24  
|||  
Db 125 AAGCCCTGTGCC 136  
|||

## RESULT 2

ACD96468/c  
ID ACD96468 standard; cDNA; 670 BP.  
XX

AC ACD96468;

DT 23-SEP-2003 (first entry)

XX Human colon cancer cell expressed cDNA #4880.

XX Open reading frame detection; genome sequencing; colon cancer;  
KW breast cancer; population genome analysis; genetic shift; cancer;  
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;  
KW agriculture; food crop genome; resistance gene; retrovirus;  
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;  
KW gene; ss.

XX Homo sapiens.

XX US2002155438-A1.

XX 24-OCT-2002.

XX 27-SEP-1999; 99US-00406117.

XX 20-NOV-1998; 98US-00196716.

XX (SIMP/) SIMPSON A J G.

XX (NETO/) NETO E D.

XX (BRENT/) BRENTANI R R.

XX Simpson AJG, Neto ED, Brentani RR;

XX WPI; 2003-182626/18.

XX Determining open reading frames of genome of an organism e.g. a human  
PT suffering from cancer involves use of single oligonucleotide primer at  
PT low stringency for preparing single-stranded cDNA from mRNA of  
PT individual.

XX Example 9; Page 701; 959pp; English.





```
PR 01-FEB-2002; 2002US-0353301P.
PR 12-FEB-2002; 2002US-0356371P.
PR 13-FEB-2002; 2002US-0356424P.
PR 13-FEB-2002; 2002US-0356531P.
PR 20-FEB-2002; 2002US-0358239P.
PR 26-FEB-2002; 2002US-0359603P.
PR 27-FEB-2002; 2002US-0359848P.
PR 27-FEB-2002; 2002US-0359860P.
PR 15-MAR-2002; 2002US-0365049P.
PR 22-MAR-2002; 2002US-0366802P.
PR 17-MAY-2002; 2002US-0381666P.
PR 18-JUN-2002; 2002US-0389531P.
PR 19-JUN-2002; 2002US-0389910P.
PR 25-JUN-2002; 2002US-0391516P.
PR 02-JUL-2002; 2002US-0393265P.
PR 07-AUG-2002; 2002US-0401825P.
PR 09-AUG-2002; 2002US-0402395P.
PR 12-AUG-2002; 2002US-0402867P.
PR 23-AUG-2002; 2002US-0405401P.
PR 23-AUG-2002; 2002US-0405820P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Alsobrook JP, Bader JS, Berghs C, Burgess CE, Casman SJ,
PI Catterton E, Chaudhuri A, Edinger SR, Ellerman K, Gerlach VL,
PI Gorman L, Guo X, Herrmann JL, Ji W, Khrantsov NV, Li L, Miller CE,
PI Ort T, Patturajan M, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA,
PI Spytek KA, Vernet CAM, Zhong H, Zhong M;
XX WPI; 2003-646149/61.
DR P-PSDB; ADM29354.
XX
XX New NOVX polypeptide, useful for the manufacture of a medicament for
PT treating e.g., cancer or immune associated disorders.
XX
XX Claim 20; SEQ ID NO 97; 606pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC human proteins (NOV proteins). The DNA and protein sequences of the
CC invention are useful for the manufacture of a medicament for treating a
CC syndrome associated with a human disease comprising a pathology
CC associated with the protein, such as: cancer or immune associated
CC disorders. The present DNA sequence encodes a NOV protein of the
CC invention.
XX
XX Sequence 1255 BP; 256 A; 415 C; 360 G; 224 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.12e-10 Length: 1255
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 11 Gaps: 0
US-09-632-036f-6 (1-24) x ADM29353 (1-1255)
Qy 1 ProteuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 971 CCCCCTGCACCAACCAAGGTTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGACG 1030
Qy 21 LysProCysAla 24
Db 1031 AAGCCCTGTGCC 1042
RESULT 5
ADF16383
ID ADF16383 standard; DNA; 1260 BP.
XX
XX ADF16383;
XX
XX 12-FEB-2004 (first entry)
XX
```

```
DE Human albumin fusion protein-related DNA sequence SeqID1475.
XX
XX albumin fusion protein; albumin activity; human serum albumin;
KW serum osmotic pressure; shelf-life; stability; antidiabetic;
KW gene therapy; diabetes mellitus; human; gene; ds.
XX
XX Homo sapiens.
XX WO2003060071-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US040891.
XX
XX 21-DEC-2001; 2001US-0341811P.
PR 24-JAN-2002; 2002US-0350358P.
PR 28-JAN-2002; 2002US-0351360P.
PR 26-FEB-2002; 2002US-0359370P.
PR 28-FEB-2002; 2002US-0360000P.
PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-MAY-2002; 2002US-0382617P.
PR 28-MAY-2002; 2002US-0383123P.
PR 05-JUN-2002; 2002US-0385708P.
PR 10-JUL-2002; 2002US-0394625P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 18-SEP-2002; 2002US-0411426P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
XX Ballance DJ, Turner AJ, Rosen CA, Haseltine WA;
PI WPI; 2003-598517/56.
XX P-PSDB; ADF16709.
DR
XX
XX New albumin fusion protein, useful for preparing a composition for
PT treating diabetes mellitus.
XX
XX Example 4; SEQ ID NO 1475; 24pp; English.
XX
XX This invention relates to a novel albumin fusion protein having albumin
CC or biological activity. Human serum albumin is responsible for a
CC significant proportion of the osmotic pressure of serum and also
CC functions as a carrier of endogenous and exogenous ligands. The fusion of
CC albumin to a therapeutic protein may increase shelf-life and stability of
CC the therapeutic protein. The albumin fusion protein of the invention may
CC allow production of compositions with antidiabetic activity whilst the
CC nucleotide sequence which encodes it may be useful for gene therapy. The
CC albumin fusion protein which is useful for preparing a composition for treating
CC diabetes mellitus. The present sequence is a DNA sequence which encodes a
CC therapeutic protein which was fused with human albumin to create a novel
CC albumin fusion protein of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/publishedpct_sequences
XX
XX Sequence 1260 BP; 238 A; 441 C; 340 G; 241 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.13e-10 Length: 1260
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
```

```
Query Match: 96.97% Indels: 0
DB: 10 Gaps: 0
US-09-632-036F-6 (1-24) x ADP16383 (1-1260)
QY 1 ProteuHisAenGlnGluValThrAlaGluAepGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1005
QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017
RESULT 6
ADZ47818
ID ADZ47818 standard; DNA; 1320 BP.
AC ADZ47818;
XX 30-JUN-2005 (first entry)
XX DNA encoding human ErbB-2 variant VI.
XX epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;
KW prognosis; cancer; cytostatic; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 151..1314
FT /tag= a
FT /product= "ErbB-2 variant VI"
FT CDS 1171..1314
FT /tag= b
FT /product= "Active region of ErbB-2 variant VI"
FT /note= "Given as SEQ ID No:27, specifically claimed in
FT Claim 170"
XX WO200503133-A2.
XX 14-APR-2005.
XX 04-OCT-2004; 2004WO-US030903.
XX 03-OCT-2003; 2003US-0507953P.
XX 10-NOV-2003; 2003US-0518321P.
XX 09-AUG-2004; 2004US-0599583P.
XX 07-SEP-2004; 2004US-0607326P.
XX (COMP-) COMPUTEN LTD.
XX (COMP-) COMPUTEN INC.
XX Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;
XX Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;
XX WPI; 2005-285403/29.
XX P-PSDB; ADZ47818.
XX New isolated polynucleotide encoding at least an active portion of an
XX ErbB-2 polypeptide, useful for diagnosing or treating diseases with
XX aberrant expression or activity of the ErbB-2 polypeptides, such as
XX cancer.
XX Claim 70; SEQ ID NO 25; 246pp; English.
XX The invention relates to the isolation of polynucleotide sequences
XX encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as
XX HER-2 or neu) polypeptides. Also described are methods, compositions, and
XX kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and
XX treatment of ErbB-2 related cancer. The sequences, methods, and
XX compositions of the present invention are useful for diagnosing,
XX prognosing, and/or treating diseases or conditions associated with
XX aberrant expression or activity of the ErbB-2 polypeptides, such as
XX cancer.
CC cancer. This sequence encodes human ErbB-2 variant VI.
XX
SQ Sequence 1320 BP; 261 A; 462 C; 371 G; 226 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.19e-10 Length: 1320
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 14 Gaps: 0
US-09-632-036F-6 (1-24) x ADZ47818 (1-1320)
QY 1 ProteuHisAenGlnGluValThrAlaGluAepGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1155
QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167
RESULT 7
ADZ69874
ID ADZ69874 standard; cDNA; 1459 BP.
XX AC ADZ69874;
XX 14-JUL-2005 (first entry)
XX p185neu oncoprotein-encoding cDNA, SEQ ID 5.
XX gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;
XX vaccine; ss; pharmaceutical.
XX Homo sapiens.
XX Rattus sp.
XX Chimeric.
XX WO2005039618-A1.
XX 06-MAY-2005.
XX 06-OCT-2004; 2004WO-EP011161.
XX 09-OCT-2003; 2003IT-MI001942.
XX (INDE-) INDENA SPA.
XX Amici A, Cavallo F, Forni G, Marchini C;
XX WPI; 2005-322913/33.
XX New DNA transfer vector containing a sequence coding for a p185neu-
XX fragment, useful for preparing a pharmaceutical composition for treating
XX tumors, and for preparing DNA vaccine.
XX Claim 1; SEQ ID NO 5; 64pp; English.
XX The invention relates to a DNA transfer vector containing a sequence
XX coding for a p185neu-fragment. Also described are: a pharmaceutical
XX composition comprising the DNA vector in admixture with vehicles and
XX excipient; and a combined pharmaceutical preparation containing at least
XX two different plasmids for simultaneous, sequential or separate
XX therapeutic use. The DNA transfer vector is a plasmid, which further
XX contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid
XX is useful in the therapeutic treatment of subjects at risk of developing
XX p185neu positive tumors, or of patients having primary tumors, metastasis
XX or relapses of p185neu-expressing tumors. The plasmid is also useful for
XX the preparation of DNA vaccine. The present sequence represents a p185neu
XX -encoding fragment used in the plasmids of the invention.
XX Sequence 1459 BP; 272 A; 467 C; 434 G; 286 T; 0 U; 0 Other;
SQ
```

```
Alignment Scores:
Pred. No.: 1.35e-10 Length: 1459
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 14 Gaps: 0

US 09-632-036F-6 (1-24) x ADZ69874 (1-1459)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 350 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 409
   |||||

QY 21 LysProCysAla 24
   |||||
Db 410 AAGCCCTGTGCC 421

RESULT 8
ACC83226
ID ACC83226 standard; DNA; 1555 BP.
XX
AC ACC83226;
XX
DT 05-SEP-2003 (first entry)
XX
DE Expression vector pCDNA3-hSIDDD DNA.
XX
KW DNA tumour; vaccine; pCDNA3-hSIDDD; vector; ds.
XX
OS Unidentified.
XX
PN CN1380107-A.
XX
PD 20-NOV-2002.
XX
PF 27-SEP-2001; 2001CN-00126890.
XX
PR 27-SEP-2001; 2001CN-00126890.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI He Q, Wang Y, Zhang Y;
DR WPI; 2003-222532/22.
XX
PT DNA tumor vaccine and its preparation method.
XX
PS Claim 1; Page 8; 20pp; Chinese.
XX
CC The invention relates to a novel DNA tumour vaccine. Also disclosed is a
CC method for preparing the vaccine comprising, adopting a specific fragment
CC hSIDDD of clonal human HER-2/neu and utilising NotI/XhoI double enzyme
CC sections to directionally insert them into the pcDNA3 plasmid to
CC construct the human HER-2/neu gene eucaryon expression vector pCDNA3-
CC hSIDDD. The current sequence represents the expression vector pCDNA3-hSIDDD
CC DNA
XX
SQ Sequence 1555 BP; 306 A; 527 C; 444 G; 278 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.46e-10 Length: 1555
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 8 Gaps: 0

US-09-632-036F-6 (1-24) x ACC83226 (1-1555)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
```

```
Db 1091 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1150
QY 21 LysProCysAla 24
   |||||
Db 1151 AAGCCCTGTGCC 1162

RESULT 9
ADZ69873
ID ADZ69873 standard; cDNA; 1699 BP.
XX
AC ADZ69873;
XX
DT 14-JUL-2005 (first entry)
XX
DE p185neu oncoprotein-encoding cDNA, SEQ ID 4.
XX
KW gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;
KW vaccine; ss; pharmaceutical.
XX
OS Homo sapiens.
OS Rattus sp.
OS Chimeric.
XX
PN WO2005039618-A1.
XX
PD 06-MAY-2005.
XX
PF 06-OCT-2004; 2004WO-EP011161.
XX
PR 09-OCT-2003; 2003IT-MI001942.
XX
PA (INDE-) INDENA SPA.
XX
PI Amici A, Cavallo F, Forni G, Marchini C;
XX
DR WPI; 2005-322913/33.
XX
PT New DNA transfer vector containing a sequence coding for a p185neu-
PT fragment, useful for preparing a pharmaceutical composition for treating
PT tumors, and for preparing DNA vaccine.
XX
PS Claim 1; SEQ ID NO 4; 64pp; English.
XX
CC The invention relates to a DNA transfer vector containing a sequence
CC coding for a p185neu-fragment. Also described are: a pharmaceutical
CC composition comprising the DNA vector in admixture with vehicles and
CC excipient; and a combined pharmaceutical preparation containing at least
CC two different plasmids for simultaneous, sequential or separate
CC therapeutic use. The DNA transfer vector is a plasmid, which further
CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid
CC is useful in the therapeutic treatment of subjects at risk of developing
CC p185neu positive tumors, or of patients having primary tumors, metastasis
CC or relapses of p185neu-expressing tumors. The plasmid is also useful for
CC the preparation of DNA vaccine. The present sequence represents a p185neu
CC -encoding fragment used in the plasmids of the invention.
XX
SQ Sequence 1699 BP; 319 A; 541 C; 505 G; 334 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.62e-10 Length: 1699
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ69873 (1-1699)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 590 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 649
   |||||

QY 21 LysProCysAla 24
```

```
Db          650 AAGCCCTGTGCC 661
RESULT 10
AAQ06828
ID AAQ06828 standard; rRNA; 1872 BP.
AC
XX AAQ06828;
XX AC
DT 06-MAR-1991 (first entry)
XX
XX Extracellular portion of the human epidermal growth factor receptor 2.
XX Human epidermal growth factor receptor 2; HER2; vaccine; cancer; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_RNA 64..96
XX FT /*tag= a
XX FT /number= 1
XX FT /note= "potential T-cell epitope"
XX FT 220..252
XX FT /*tag= b
XX FT /number= 2
XX FT /note= "potential T-cell epitope"
XX FT 337..390
XX FT /*tag= c
XX FT /number= 3
XX FT /note= "potential T-cell epitope"
XX FT 559..591
XX FT /*tag= d
XX FT /number= 4
XX FT /note= "potential T-cell epitope"
XX FT 913..945
XX FT /*tag= e
XX FT /number= 5
XX FT /note= "potential T-cell epitope"
XX FT 979..1059
XX FT /*tag= f
XX FT /number= 6
XX FT /note= "potential T-cell epitope"
XX FT 1123..1155
XX FT /*tag= g
XX FT /number= 7
XX FT /note= "potential T-cell epitope"
XX FT 1192..1224
XX FT /*tag= h
XX FT /number= 8
XX FT /note= "potential T-cell epitope"
XX FT 1297..1329
XX FT /*tag= i
XX FT /number= 9
XX FT /note= "potential T-cell epitope"
XX FT 1555..1587
XX FT /*tag= j
XX FT /number= 10
XX FT /note= "potential T-cell epitope"
XX
XX WO9014357-A.
XX
XX 29-NOV-1990.
XX
XX 10 MAY 1999; 89US 00354319.
XX
XX 19-MAY-1989; 89US-00354319.
XX
XX (GETH ) GENENTECH INC.
XX
XX Hudziak RM, Shepard HM, Ullrich A;
XX WPI; 1990-375946/50.
XX P-PSDB; AAR08222.
XX

XX HER2 extracellular domain used as vaccine - comprises sequence of at
XX least 9 aminoacid(s) prepd. using expression vector of DNA isolated from
XX human epidermal growth factor receptor.
XX
XX Disclosure; Fig 13; 49pp; English.
XX
XX This claimed sequence terminates 1 bp, pref. 24 bps, upstream of the DNA
XX portion encoding the transmembrane domain of the HER2 mol. It encodes
XX none of the transmembrane or intracellular portions of the HER2 mol. The
XX protein is antigenic in animals. The sequence can be used to produce an
XX extracellular portion of the HER2 mol. comprising at least 9 amino acids.
XX The product opens several possibilities for further research and a broad
XX range of potential therapeutic applications, eg. for the treatment of
XX mammary gland adenocarcinoma and other cancers
XX
XX Sequence 1872 BP; 377 A; 602 C; 541 G; 0 T; 352 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.83e-10 Length: 1872
XX Score: 128.00 Matches: 23
XX Percent Similarity: 95.83% Conservative: 0
XX Best Local Similarity: 95.83% Mismatches: 1
XX Query Match: 96.97% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-632-036F-6 (1-24) x AAQ06828 (1-1872)
XX
XX QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
XX |||||
XX Db 883 CCCUGCACAACCAAGAGGUGACAGCAGAGGAUGGAACACACAGCGGUGAGAGAGCAGC 942
XX
XX QY 21 LysProCysAla 24
XX |||||
XX Db 943 AAGCCCGUGGCC 954
XX
XX RESULT 11
XX ADZ69872
XX ID ADZ69872 standard; cDNA; 1939 BP.
XX
XX AC ADZ69872;
XX
XX 14-JUL-2005 (first entry)
XX
XX DE p185neu oncoprotein-encoding cDNA, SEQ ID 3.
XX
XX KW gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;
XX KW vaccine; ss; pharmaceutical.
XX
XX OS Homo sapiens.
XX OS Rattus sp.
XX OS Chimeric.
XX
XX WO2005039618-A1.
XX
XX 06-MAY-2005.
XX
XX 06-OCT-2004; 2004WO-EP011161.
XX
XX 09-OCT-2003; 2003IT-MI001942.
XX
XX (INDE-) INDENA SPA.
XX
XX Amici A, Cavallo F, Forni G, Marchini C;
XX WPI; 2005-322913/33.
XX
XX New DNA transfer vector containing a sequence coding for a p185neu-
XX fragment, useful for preparing a pharmaceutical composition for treating
XX tumors, and for preparing DNA vaccine.
XX
XX Claim 1; SEQ ID NO 3; 64pp; English.
XX
```

CC The invention relates to a DNA transfer vector containing a sequence  
CC coding for a p185neu-fragment. Also described are: a pharmaceutical  
CC composition comprising the DNA vector in admixture with vehicles and  
CC excipient; and a combined pharmaceutical preparation containing at least  
CC two different plasmids for simultaneous, sequential or separate  
CC therapeutic use. The DNA transfer vector is a plasmid, which further  
CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid  
CC is useful in the therapeutic treatment of subjects at risk of developing  
CC p185neu positive tumors, or of patients having primary tumors, metastasis  
CC or relapses of p185neu-expressing tumors. The plasmid is also useful for  
CC the preparation of DNA vaccine. The present sequence represents a p185neu  
CC -encoding fragment used in the plasmids of the invention.  
XX  
SQ Sequence 1939 BP; 373 A; 615 C; 574 G; 377 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.91e-10 Length: 1939  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ69872 (1-1939)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
ADZ47796  
ID ADZ47796 standard; DNA; 1944 BP.  
XX  
AC ADZ47796;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE DNA encoding human ErbB-2 variant II.  
XX  
KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic; gene, ds.  
XX  
OS Homo sapiens.

Key Location/Qualifiers  
FT CDS 151..1878  
FT /tag= a  
FT /product= "ErbB-2 variant II"  
FT 1864..1944  
FT /tag= b  
FT /product= "Active region of ErbB-2 variant II"  
FT /note= "Given as SEQ ID No:8, specifically claimed in  
FT Claim 10"  
XX  
PN WO2005033133-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 04-OCT-2004; 2004WO-US030903.  
XX  
PR 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX  
PA (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;

PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
DR WPI; 2005-285403/29.  
DR P-PSDB; ADZ47797.

XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
PT cancer.

XX Claim 8; SEQ ID NO 3; 246pp; English.

XX The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
CC cancer. This sequence encodes human ErbB-2 variant II.

XX Sequence 1944 BP; 370 A; 671 C; 539 G; 364 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.92e-10 Length: 1944  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ47796 (1-1944)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
ADZ57171  
ID ADZ57171 standard; cDNA; 1956 BP.  
XX  
AC ADZ57171;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Her-2/neu gene lacking intracellular region and transmembrane domain.  
XX  
KW Her-2/neu; vaccine; cancer; glycoprotein D; cytostatic; human;  
KW gene; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO2004007734-A1.  
XX  
PD 22-JAN-2004.  
XX  
PF 15-JUL-2003; 2003WO-KR001400.  
XX  
PR 16-JUL-2002; 2002KR-00041764.  
PR 12-JUN-2003; 2003KR-00038012.  
XX  
PA (PANG-) PANGENOMICS CO LTD.  
XX  
PI Lee JY, Kim D, Chung Y, Chang S, Lee K, Kang C;  
XX  
DR WPI; 2004-122962/12.  
XX

PT New Her-2/neu plasmid construct having anti-cancer activity, useful for  
 PT preparing a DNA vaccine for preventing and/or treating cancer.  
 XX  
 PS Claim 5; SEQ ID NO 3; 70pp; English.  
 XX  
 CC The invention relates to an Her-2/neu plasmid construct having anti-  
 CC cancer activity that is prepared by inserting a truncated human Her-2/neu  
 CC gene lacking the intracellular domain into plasmid pTV2 or pCK. Aslo  
 CC provided are a DNA vaccine for preventing and/or treating cancer  
 CC comprising the plasmid construct and a carrier; and a method for  
 CC preventing and/or treating cancer by administering the DNA vaccine cited  
 CC above. The construct is pHER2 (KCCM-10393), pCKTM (KCCM-10396), pNeuECD  
 CC (KCCM-10394) or pCKECD (KCCM-10395). The truncated human Her-2/neu gene  
 CC further lacks the transmembrane domain. The signal peptide of the human  
 CC Her-2/neu gene is replaced by the signal peptide of herpes simplex type 1  
 CC glycoprotein D (gD). The plasmid construct is preferably pNeuTM-gs. The  
 CC Her-2/neu gene. The cytokine gene is selected from granulocyte-macrophage  
 CC colony-stimulating factor (GM-CSF), FMS-like tyrosine kinase 3 ligand  
 CC (Flt3L), early T lymphocyte activation-1 (Eta-1), interleukin-12 (IL-12),  
 CC IL-15 and IL-18. The DNA vaccine further comprises a cytokine gene  
 CC expressing plasmid. The Her-2/neu plasmid construct is useful for  
 CC preparing a DNA vaccine for treating and/or preventing cancer. The  
 CC present sequence represents the nucleotide sequence of the human Her-  
 CC 2/neu gene lacking the intracellular region and transmembrane domain.  
 XX  
 SQ Sequence 1956 BP; 383 A; 640 C; 567 G; 366 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.93e-10 Length: 1956  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 12 Gaps: 0

US-09-632-036F-6 (1-24) x ADJ5171 (1-1956)  
 QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 Db 946 CCCTGCAACACCCAGAGGTGACAGCAGAGATGGAAACACAGCGGTGTGAGAAAGTGCAGC 1005

QY 21 LysProCysAla 24  
 Db 1006 AAGCCCTGTGCC 1017

RESULT 14  
 ADW87396  
 ID ADW87396 standard; DNA; 2028 BP.

XX AC ADW87396;

XX DT 21-APR-2005 (first entry)

XX DE Human codon-optimized HER2ECDTM DNA.

XX KW human epidermal growth factor 2; HER2ECDTM; cancer; neoplasm; Cytostatic;  
 XX vaccine; ds; gene.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..2028

XX FT /\*tag= a

XX FT /product= "HER2ECDTM"

XX PN WO2005012527-A1.

XX PD 10-FEB-2005.

XX PF 20-JUL-2004; 2004WO-EP008234.

XX PR 21-JUL-2003; 2003US-0489237P.

XX PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX PI Gallo P, Monaci P, Nuzzo M;

XX WPI; 2005-123289/13.

XX DR P-PSDB; ADW87401.

XX PT New synthetic nucleic acid molecule encoding human epidermal growth  
 PT factor 2 (HER2)/neu or HER2ECDTM protein, useful for preventing or  
 PT treating HER2-associated cancer.

XX PS Claim 7; SEQ ID NO 9; 62pp; English.

XX CC The invention relates to a synthetic nucleic acid molecule which  
 CC comprises a sequence of nucleotides: (a) encoding a human epidermal  
 CC growth factor 2 (HER2)/neu protein; (b) encoding a human HER2ECDTM  
 CC protein; and (c) encoding a variant human HER2/neu or HER2ECDTM  
 CC polypeptide. The sequences, vector, vaccine, composition, and methods are  
 CC useful for preventing or treating HER2-associated cancer. The present  
 CC sequence represents the human codon-optimized HER2ECDTM DNA.

XX SQ Sequence 2028 BP; 338 A; 802 C; 619 G; 269 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2.02e-10 Length: 2028  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADW87396 (1-2028)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 Db 946 CCCTGCAACACCCAGAGGTGACCGCGAGGACCGCACCCAGCGCTCGAGAAAGTGCAGC 1005

QY 21 LysProCysAla 24

Db 1006 AAGCCCTGTGCC 1017

RESULT 15

ADW87397

ID ADW87397 standard; DNA; 2028 BP.

XX AC ADW87397;

XX DT 21-APR-2005 (first entry)

XX DE Human wild type HER2ECDTM DNA.

XX KW human epidermal growth factor 2; HER2ECDTM; cancer; neoplasm; Cytostatic;  
 XX vaccine; ds; gene.

XX OS Homo sapiens.

XX PN WO2005012527-A1.

XX PD 10-FEB-2005.

XX PF 20-JUL-2004; 2004WO-EP008234.

XX PR 21-JUL-2003; 2003US-0489237P.

XX PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX PI Gallo P, Monaci P, Nuzzo M;

XX WPI; 2005-123289/13.

XX PT New synthetic nucleic acid molecule encoding human epidermal growth  
 PT factor 2 (HER2)/neu or HER2ECDTM protein, useful for preventing or

PT treating HER2-associated cancer.  
XX Disclosure; SEQ ID NO 10; 62pp; English.  
PS  
XX The invention relates to a synthetic nucleic acid molecule which  
CC comprises a sequence of nucleotides: (a) encoding a human epidermal  
CC growth factor 2 (HER2)/neu protein; (b) encoding a human HER2ECDTM  
CC polypeptide. The sequences, vector, vaccine, composition, and methods are  
CC useful for preventing or treating HER2-associated cancer. The present  
CC sequence represents the human wild-type HER2ECDTM DNA.  
XX  
SQ Sequence 2028 BP; 389 A; 656 C; 591 G; 392 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2,028-10 Length: 2028  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0  
US-09-632-036F-6 (1-24) x ADW87397 (1-2028)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluIysCysSer 20  
Db 946 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCACG 1005  
QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 16  
ADJ57170  
ID ADJ57170 standard; cDNA; 2052 BP.  
XX  
AC ADJ57170;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human Her-2/neu gene lacking intracellular region.  
XX  
KW Her-2/neu; vaccine; cancer; glycoprotein D; cytokine; cytostatic; human;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004007734-A1.  
XX  
PD 22-JAN-2004.  
XX  
PF 15-JUL-2003; 2003WO-KR001400.  
XX  
PR 16-JUL-2002; 2002KR-00041764.  
XX  
XX 16-JUN-2003; 2003KR-00038012.  
XX  
PA (PANG-) PANGENOMICS CO LTD.  
XX  
PI Lee JY, Kim D, Chung Y, Chang S, Lee K, Kang C;  
XX  
XX WPI; 2004-122962/12.  
XX  
PT New Her-2/neu plasmid construct having anti-cancer activity, useful for  
PT preparing a DNA vaccine for preventing and/or treating cancer.  
XX  
XX Claim 2; SEQ ID NO 2; 70pp; English.  
XX  
XX The invention relates to an Her-2/neu plasmid construct having anti-  
CC cancer activity that is prepared by inserting a truncated human Her-2/neu  
CC gene lacking the intracellular domain into plasmid pTV2 or pCK. Aslo  
CC provided are a DNA vaccine for preventing and/or treating cancer  
CC comprising the plasmid construct and a carrier; and a method for  
CC preventing and/or treating cancer by administering the DNA vaccine cited

CC above. The construct is pNeuTM (KCCM-10393), pCKTM (KCCM-10396), pNeuECD  
CC (KCCM-10394) or pCKECD (KCCM-10395). The truncated human Her-2/neu gene  
CC further lacks the transmembrane domain. The signal peptide of the human  
CC Her-2/neu gene is replaced by the signal peptide of herpes simplex type 1  
CC glycoprotein D (gp). The plasmid construct is preferably pNeuTM-gps. The  
CC plasmid construct further translates a cytokine gene besides the human  
CC Her-2/neu gene. The cytokine gene is selected from granulocyte-macrophage  
CC colony-stimulating factor (GM-CSF), FMS-like tyrosine kinase 3 ligand  
CC (Flt3L), early T lymphocyte activation-1 (Eta-1), interleukin-12 (IL-12),  
CC IL-15 and IL-18. The DNA vaccine further comprises a cytokine gene  
CC expressing plasmid. The Her-2/neu plasmid construct is useful for  
CC preparing a DNA vaccine for treating and/or preventing cancer. The  
CC present sequence represents the nucleotide sequence of the human Her-  
CC 2/neu gene lacking the intracellular region.  
XX  
SQ Sequence 2052 BP; 397 A; 662 C; 601 G; 392 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2,052-10 Length: 2052  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 12 Gaps: 0  
US-09-632-036F-6 (1-24) x ADJ57170 (1-2052)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluIysCysSer 20  
Db 946 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCACG 1005  
QY 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017  
RESULT 17  
ADJ47808  
ID ADJ47808 standard; DNA; 2061 BP.  
XX  
AC ADJ47808;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE DNA encoding human ErbB-2 variant V.  
XX  
KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 151..1995  
FT /\*tag= a  
FT /product= "ErbB-2 variant V"  
XX  
PN WO2005033133-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 04-OCT-2004; 2004WO-US030903.  
XX  
PR 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX  
PA (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.  
XX  
XX Shesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
XX WPI; 2005-285403/29.  
DR



DR P-PSDB; AD247809.

XX New isolated polynucleotide encoding at least an active portion of an

PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with

PT aberrant expression or activity of the ErbB-2 polypeptides, such as

PT cancer.

XX Claim 8; SEQ ID NO 15; 246pp; English.

XX The invention relates to the isolation of polynucleotide sequences

CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as

CC HER-2 or neu) polypeptides. Also described are methods, compositions, and

CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and

CC treatment of ErbB-2 related cancer. The sequences, methods, and

CC compositions of the present invention are useful for diagnosing,

CC prognosing, and/or treating diseases or conditions associated with

CC aberrant expression or activity of the ErbB-2 polypeptides, such as

CC cancer. This sequence encodes human ErbB-2 variant V.

XX

SQ Sequence 2061 BP; 392 A; 706 C; 568 G; 395 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,06e-10	Length:	2061
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	14	Gaps:	0

US-09-632-036F-6 (1-24) x AD247808 (1-2061)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20

Db 1096 CCCCTGCACACCAAGAGGTGCACAGCAGAGTGGACACAGCGGTGTGAGAGTGCAGC 1155

QY 21 LysProCysAla 24

Db 1156 AAGCCCTGTGCC 1167

RESULT 18

AD269880

ID AD269880 standard; cDNA; 2086 BP.

AC AD269880;

XX 14-JUL-2005 (first entry)

XX p185neu oncoprotein-encoding cDNA, SEQ ID 11.

XX gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;

KW vaccine; ss; pharmaceutical.

XX Homo sapiens.

OS Rattus sp.

OS Chimeric.

XX WO2005039618-A1.

XX 06-MAY-2005.

XX 06-OCT-2004; 2004WO-EP011161.

XX 09-OCT-2003; 2003IT-MI001942.

XX (INDE-) INDENA SPA.

XX Amici A, Cavallo F, Fornì G, Marchini C;

XX WPI; 2005-322913/33.

XX New DNA transfer vector containing a sequence coding for a p185neu-

PT fragment, useful for preparing a pharmaceutical composition for treating

PT tumors, and for preparing DNA vaccine.

XX Claim 1; SEQ ID NO 11; 64pp; English.

XX The invention relates to a DNA transfer vector containing a sequence

CC coding for a p185neu-fragment. Also described are: a pharmaceutical

CC composition comprising the DNA vector in admixture with vehicles and

CC excipient; and a combined pharmaceutical preparation containing at least

CC two different plasmids for simultaneous, sequential or separate

CC therapeutic use. The DNA transfer vector is a plasmid, which further

CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid

CC is useful in the therapeutic treatment of subjects at risk of developing

CC p185neu positive tumors, or of patients having primary tumors, metastasis

CC or relapses of p185neu-expressing tumors. The plasmid is also useful for

CC the preparation of DNA vaccine. The present sequence represents a p185neu

CC -encoding fragment used in the plasmids of the invention.

XX

SQ Sequence 2086 BP; 405 A; 565 C; 612 G; 404 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,09e-10	Length:	2086
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	14	Gaps:	0

US-09-632-036F-6 (1-24) x AD269880 (1-2086)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20

Db 977 CCCCTGCACACCAAGAGGTGCACAGCAGAGTGGACACACAGCGGTGTGAGAGTGCAGC 1036

QY 21 LysProCysAla 24

Db 1037 AAGCCCTGTGCC 1048

RESULT 19

AD269879

ID AD269879 standard; cDNA; 2086 BP.

XX AD269879;

XX 14-JUL-2005 (first entry)

XX p185neu oncoprotein-encoding cDNA, SEQ ID 10.

XX gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;

KW vaccine; ss; pharmaceutical.

XX Homo sapiens.

OS Rattus sp.

OS Chimeric.

XX WO2005039618-A1.

XX 06-MAY-2005.

XX 06-OCT-2004; 2004WO-EP011161.

XX 09-OCT-2003; 2003IT-MI001942.

XX (INDE-) INDENA SPA.

XX Amici A, Cavallo F, Fornì G, Marchini C;

XX WPI; 2005-322913/33.

XX New DNA transfer vector containing a sequence coding for a p185neu-

PT fragment, useful for preparing a pharmaceutical composition for treating

PT tumors, and for preparing DNA vaccine.

XX Claim 1; SEQ ID NO 10; 64pp; English.

XX

CC The invention relates to a DNA transfer vector containing a sequence  
CC coding for a p185neu-fragment. Also described are: a pharmaceutical  
CC composition comprising the DNA vector in admixture with vehicles and  
CC excipient; and a combined pharmaceutical preparation containing at least  
CC two different plasmids for simultaneous, sequential or separate  
CC therapeutic use. The DNA transfer vector is a plasmid, which further  
CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid  
CC is useful in the therapeutic treatment of subjects at risk of developing  
CC p185neu positive tumors, or of patients having primary tumors, metastasis  
CC or relapses of p185neu-expressing tumors. The plasmid is also useful for  
CC the preparation of DNA vaccine. The present sequence represents a p185neu  
CC -encoding fragment used in the plasmids of the invention.

XX  
SQ Sequence 2086 BP; 403 A; 664 C; 615 G; 404 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.09e-10 Length: 2086  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ69879 (1-2086)

OY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 977 CCCCTGCACAAACGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1036

OY 21 LysProCysAla 24

Db 1037 AAGCCCTGTGCC 1048

RESULT 20

ADZ69881  
ID ADZ69881 standard; cDNA; 2086 BP.

XX AC ADZ69881;

XX 14-JUL-2005 (first entry)

XX p185neu oncoprotein-encoding cDNA, SEQ ID 12.

XX gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;  
KW vaccine; ss; pharmaceutical.

XX Homo sapiens.

OS Rattus sp.

OS Chimeric.

XX WO2005039618-A1.

XX 06-MAY-2005.

XX 06-JUL-2004; 2004WO-EP011161.

XX 09-OCT-2003; 2003IT-MI001942.

XX (INDE-) INDENA SPA.

XX Amici A, Cavallo F, Forni G, Marchini C;

XX WPI; 2005-322913/33.

XX New DNA transfer vector containing a sequence coding for a p185neu-  
PT fragment, useful for preparing a pharmaceutical composition for treating  
PT tumors, and for preparing DNA vaccine.

XX Claim 1; SEQ ID NO 12; 64pp; English.

XX The invention relates to a DNA transfer vector containing a sequence  
CC coding for a p185neu-fragment. Also described are: a pharmaceutical  
CC composition comprising the DNA vector in admixture with vehicles and

CC excipient; and a combined pharmaceutical preparation containing at least  
CC two different plasmids for simultaneous, sequential or separate  
CC therapeutic use. The DNA transfer vector is a plasmid, which further  
CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid  
CC is useful in the therapeutic treatment of subjects at risk of developing  
CC p185neu positive tumors, or of patients having primary tumors, metastasis  
CC or relapses of p185neu-expressing tumors. The plasmid is also useful for  
CC the preparation of DNA vaccine. The present sequence represents a p185neu  
CC -encoding fragment used in the plasmids of the invention.

XX  
SQ Sequence 2086 BP; 410 A; 656 C; 609 G; 411 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.09e-10 Length: 2086  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ69881 (1-2086)

OY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 977 CCCCTGCACAAACGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1036

OY 21 LysProCysAla 24

Db 1037 AAGCCCTGTGCC 1048

RESULT 21

ADZ69882

ID ADZ69882 standard; cDNA; 2086 BP.

XX AC ADZ69882;

XX 14-JUL-2005 (first entry)

XX p185neu oncoprotein-encoding cDNA, SEQ ID 13.

XX gene transfer; vector; plasmid; p185neu; cancer; neoplasm; cytostatic;  
KW vaccine; ss; pharmaceutical.

XX Homo sapiens.

OS Rattus sp.

OS Chimeric.

XX WO2005039618-A1.

XX 06-MAY-2005.

XX 06-OCT-2004; 2004WO-EP011161.

XX 09-OCT-2003; 2003IT-MI001942.

XX (INDE-) INDENA SPA.

XX Amici A, Cavallo F, Forni G, Marchini C;

XX WPI; 2005-322913/33.

XX New DNA transfer vector containing a sequence coding for a p185neu-  
PT fragment, useful for preparing a pharmaceutical composition for treating  
PT tumors, and for preparing DNA vaccine.

XX Claim 1; SEQ ID NO 13; 64pp; English.

XX The invention relates to a DNA transfer vector containing a sequence  
CC coding for a p185neu-fragment. Also described are: a pharmaceutical  
CC composition comprising the DNA vector in admixture with vehicles and  
CC excipient; and a combined pharmaceutical preparation containing at least  
CC two different plasmids for simultaneous, sequential or separate  
CC therapeutic use. The DNA transfer vector is a plasmid, which further

CC contains a CMV transcription promoter and 4 or 8 CpG motifs. The plasmid  
CC is useful in the therapeutic treatment of subjects at risk of developing  
CC p185neu positive tumors, or of patients having primary tumors, metastasis  
CC or relapse of p185neu-expressing tumors. The plasmid is also useful for  
CC the preparation of DNA vaccine. The present sequence represents a p185neu  
CC -encoding fragment used in the plasmids of the invention.

XX  
SQ Sequence 2086 BP; 413 A; 654 C; 608 G; 411 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	2,09e-10	Length:	2086
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	14	Gaps:	0

US-09-632-036F-6 (1-24) x ADZ69882 (1-2086)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 977 CCCCTGCACACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1036

Qy 21 LysProCysAla 24

Db 1037 AAGCCTGTGCC 1048

## RESULT 22

ADG17937  
ID ADG17937 standard; DNA; 2132 BP.

XX AC ADG17937;

XX DT 26-FEB-2004 (first entry)

XX DE Human HER2-sv form 97 gene encoding sequence.

XX KW Cytostatic; Gene therapy; HER-2sv polypeptide-related disease; cancer;  
XX ds.

XX OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	78..2132
FT	/*tag= a

XX WO2003087338-A2.

XX PD 23-OCT-2003.

XX PF 11-APR-2003; 2003WO-US011392.

XX PR 11-APR-2002; 2002US-0371912P.

XX PA (AMGE-) AMGEN INC.

XX PI Tatarewicz S, Jing S;

XX DR WPI; 2003-903161/82.

XX DR P-PSDB; ADG17938.

XX New-HER-2 splice variant, HER-2sv, nucleic acid molecules, useful for  
PT treating, preventing or ameliorating a HER-2sv polypeptide-related  
PT disease, condition or disorder, e.g. breast, ovarian, gastric, lung or  
PT oral cancer.

XX PS Claim 1; SEQ ID NO 3; 141pp; English.

XX The present invention relates to a new isolated nucleic acid molecule.  
CC The nucleic acid molecule, polypeptide, agent, composition and methods  
CC are useful for treating, preventing or ameliorating a HER-2sv polypeptide  
CC -related disease, condition or disorder, e.g. cancer, such as breast,  
CC ovarian, gastric, lung or oral cancer. The present sequence represents

CC human HER2-sv form 97 gene encoding sequence.  
XX  
SQ Sequence 2132 BP; 408 A; 692 C; 622 G; 410 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	2,15e-10	Length:	2132
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	10	Gaps:	0

US-09-632-036F-6 (1-24) x ADG17937 (1-2132)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1023 CCCCTGCACACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1082

Qy 21 LysProCysAla 24

Db 1083 AAGCCTGTGCC 1094

## RESULT 23

ADG17943

ID ADG17943 standard; DNA; 2149 BP.

XX AC ADG17943;

XX DT 26-FEB-2004 (first entry)

XX DE Human HER2-sv form 184 gene encoding sequence.

XX KW Cytostatic; Gene therapy; HER-2sv polypeptide-related disease; cancer;  
XX ds.

XX OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	1..2145
FT	/*tag= a

XX WO2003087338-A2.

XX PD 23-OCT-2003.

XX PF 11-APR-2003; 2003WO-US011392.

XX PR 11-APR-2002; 2002US-0371912P.

XX PA (AMGE-) AMGEN INC.

XX PI Tatarewicz S, Jing S;

XX DR WPI; 2003-903161/82.

XX DR P-PSDB; ADG17944.

XX New HER-2 splice variant, HER-2sv, nucleic acid molecules, useful for  
PT treating, preventing or ameliorating a HER-2sv polypeptide-related  
PT disease, condition or disorder, e.g. breast, ovarian, gastric, lung or  
PT oral cancer.

XX PS Claim 1; SEQ ID NO 9; 141pp; English.

XX The present invention relates to a new isolated nucleic acid molecule.  
CC The nucleic acid molecule, polypeptide, agent, composition and methods  
CC are useful for treating, preventing or ameliorating a HER-2sv polypeptide  
CC -related disease, condition or disorder, e.g. cancer, such as breast,  
CC ovarian, gastric, lung or oral cancer. The present sequence represents  
CC human HER2-sv form 184 gene encoding sequence.

XX SQ Sequence 2149 BP; 414 A; 688 C; 629 G; 418 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2.17e-10 Length: 2149  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
Gaps: 0

US-09-632-036F-6 (1-24) x ADG17943 (1-2149)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluIlysCysSer 20  
|||  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGTGGAAACACGCGGTGTGAGAAAGTCAGC 1005

QY 21 LysProCysAla 24  
|||  
Db 1006 AAGCCCTGTGCC 1017

RESULT 24  
ADG17939  
ID ADG17939 standard; DNA; 2164 BP.  
XX  
AC ADG17939;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Human HER2-sv form 119 gene encoding sequence.  
XX  
KW Cytostatic; Gene therapy; HER-2sv polypeptide-related disease; cancer;  
KW ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2164  
FT /\*tag= a  
XX  
PN WO2003087338-A2.  
XX  
PD 23-OCT-2003.  
XX  
PF 11-APR-2003; 2003WO-US011392.  
XX  
PR 11-APR-2002; 2002US-0371912P.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Tatarewicz S, Jing S;  
XX  
DR WPI; 2003-903161/82.  
DR P-PSDB; ADG17940.  
XX  
PT New HER-2 splice variant, HER-2sv, nucleic acid molecules, useful for  
PT treating, preventing or ameliorating a HER-2sv polypeptide-related  
PT disease, condition or disorder, e.g. breast, ovarian, gastric, lung or  
PT oral cancer.  
XX  
PS Claim 1; SEQ ID NO 5; 141pp; English.  
XX  
CC The present invention relates to a new isolated nucleic acid molecule.  
CC The nucleic acid molecule, polypeptide, agent, composition and methods  
CC are useful for treating, preventing or ameliorating a HER-2sv polypeptide  
CC -related disease, condition or disorder, e.g. cancer, such as breast,  
CC ovarian, gastric, lung or oral cancer. The present sequence represents  
CC human HER2-sv form 119 gene encoding sequence.  
XX  
SQ Sequence 2164 BP; 417 A; 697 C; 625 G; 425 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.19e-10 Length: 2164  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0

DB: 10 Gaps: 0

US-09-632-036F-6 (1-24) x ADG17939 (1-2164)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluIlysCysSer 20  
|||  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGTGGAAACACGCGGTGTGAGAAAGTCAGC 1005

QY 21 LysProCysAla 24  
|||  
Db 1006 AAGCCCTGTGCC 1017

RESULT 25  
ADM29355  
ID ADM29355 standard; DNA; 2261 BP.  
XX  
AC ADM29355;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human novel protein NOV20c coding sequence.  
XX  
KW human; novel protein; NOV; cancer; immune associated disorder; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2003064628-A2.  
XX  
PD 07-AUG-2003.  
XX  
PF 03-FEB-2003; 2003WO-US003401.  
XX  
PR 01-FEB-2002; 2002US-0353287P.  
PR 01-FEB-2002; 2002US-0353301P.  
PR 12-FEB-2002; 2002US-0356371P.  
PR 12-FEB-2002; 2002US-0356424P.  
PR 13-FEB-2002; 2002US-0356531P.  
PR 20-FEB-2002; 2002US-0358239P.  
PR 26-FEB-2002; 2002US-0359603P.  
PR 27-FEB-2002; 2002US-0359848P.  
PR 27-FEB-2002; 2002US-0359860P.  
PR 15-MAR-2002; 2002US-0365049P.  
PR 22-MAR-2002; 2002US-0366802P.  
PR 17-MAY-2002; 2002US-0381666P.  
PR 18-JUN-2002; 2002US-0389531P.  
PR 19-JUN-2002; 2002US-0389910P.  
PR 25-JUN-2002; 2002US-0391516P.  
PR 02-JUL-2002; 2002US-0393265P.  
PR 07-AUG-2002; 2002US-0401825P.  
PR 09-AUG-2002; 2002US-0402395P.  
PR 12-AUG-2002; 2002US-0402867P.  
PR 23-AUG-2002; 2002US-0405401P.  
PR 23-AUG-2002; 2002US-0405820P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Alsobrook JP, Bader JS, Berghs C, Burgess CE, Casman SJ;  
PI Catterton E, Chaudhuri A, Edinger SR, Ellerman K, Gerlach VL;  
PI Gorman L, Guo X, Herrmann JL, Khramtsov NV, Li L, Miller CE;  
PI Ort T, Patturajan M, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;  
PI Spytek KA, Vernet CAM, Zhong H, Zhong M;  
XX  
DR WPI; 2003-646149/61.  
DR P-PSDB; ADM29356.  
XX  
PT New NOVX polypeptide, useful for the manufacture of a medicament for  
PT treating e.g., cancer or immune associated disorders.  
XX  
PS Claim 20; SEQ ID NO 99; 606pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of novel  
CC human proteins (NOV proteins). The DNA and protein sequences of the  
CC invention are useful for the manufacture of a medicament for treating a

CC syndrome associated with a human disease comprising a pathology  
CC associated with the protein, such as: cancer or immune associated  
CC disorders. The present DNA sequence encodes a NOV protein of the  
CC invention.

XX  
SQ Sequence 2261 BP; 451 A; 711 C; 673 G; 426 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,31e-10 Length: 2261  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservatives: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x ADM29355 (1-2261)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 961 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1020  
|||

Qy 21 LysProCysAla 24  
|||  
Db 1021 AAGCCCTGTGCC 1032  
|||

RESULT 26  
ADZ47794  
ID ADZ47794 standard; DNA; 2320 BP.  
XX  
AC ADZ47794;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE DNA encoding human ErbB-2 variant I.  
XX  
KW epidermal growth factor receptor-2; ErbB-2; HER-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic; gene; ds.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FH 151..2187  
FT CDS /\*tag= a  
FT /product= "ErbB-2 variant I"  
FT CDS 2097..2320  
FT /\*tag= b  
FT /product= "Active region of ErbB-2 variant I"  
FT /note= "Given as SEQ ID No:7, specifically claimed in  
FT Claim 5"  
XX

PN WO200503133-A2.  
XX  
XX  
PD 14-APR-2005.  
XX  
XX  
PF 04-OCT-2004; 2004WO-US030903.  
XX  
PR 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX  
XX (COMP-) COMPUEN LTD.  
PA (COMP-) COMPUEN INC.  
XX  
XX Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
XX WPI; 2005-285403/29.  
DR P-PSDB; ADZ47795.  
XX  
XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as

PT cancer.  
PS Claim 3; SEQ ID NO 1; 246pp; English.  
XX  
CC The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
CC cancer. This sequence encodes human ErbB-2 variant I.

XX  
SQ Sequence 2320 BP; 448 A; 771 C; 673 G; 428 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,38e-10 Length: 2320  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservatives: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0

US-09-632-036F-6 (1-24) x ADZ47794 (1-2320)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 1096 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155  
|||

Qy 21 LysProCysAla 24  
|||  
Db 1156 AAGCCCTGTGCC 1167  
|||

RESULT 27  
AAT72725  
ID AAT72725 standard; cDNA; 2385 BP.  
XX  
AC AAT72725;  
XX  
DT 17-SEP-1997 (first entry)  
XX  
DE Her2-GM-CSF immunostimulant fusion protein DNA.  
XX  
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;  
KW growth factor receptor; oncogene; immunostimulant; cancer; therapy; ss.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FH 11..2359  
FT CDS /\*tag= a  
FT /product= "GM-CSF-Her2 fusion protein"  
FT mRNA 11..1969  
FT /\*tag= b  
FT /product= "Her2"  
FT mRNA 1970..1975  
FT /\*tag= c  
FT /product= "Leu-Glu linker"  
FT mRNA 1976..2359  
FT /\*tag= d  
FT /product= "GM-CSF"  
XX  
XX WO9724438-A1.  
XX  
XX 10-JUL-1997.  
XX  
XX 23-DEC-1996; 96WO-US020241.  
XX  
XX 28-DEC-1995; 95US-00579823.  
XX  
XX (ACTI-) ACTIVATED CELL THERAPY INC.

PI Laus R, Ruegg CL, Wu H;  
XX WPI; 1997-363674/33.  
DR P-PSDB; AAW19764.  
XX  
XX Potent APC that activates T-cells to give multivalent cellular immune  
PT response - can also induce a cytotoxic T-cell response in a vertebrate  
PT subject.  
XX  
XX Disclosure; Fig 8; 45pp; English.  
PS  
XX A nucleic acid molecule (AAT72725) codes for a fusion protein (AAW19764)  
CC comprising granulocyte-macrophage colony stimulating factor (GM-CSF) and  
CC Her2, a growth factor receptor that is over-expressed in breast and  
CC ovarian cancer cells. It was prep'd. by PCR amplification of Her2 cDNA  
CC from a breast cancer cell line and fusion to GM-CSF cDNA. Fusion  
CC expression vectors can be used to transfect mammalian and insect cells.  
CC The Her2-GM-CSF fusion protein is used to generate anti-Her2 immunity.  
CC Tumour cells are eliminated by cytotoxic T lymphocytes activated in vivo  
CC or in vitro by exposure to antigen-presenting cells exposed to the fusion  
CC protein  
XX  
SQ Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 2,46e-10 Length: 2385  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 2 Gaps: 0  
  
US-09-632-036F-6 (1-24) x AAT72725 (1-2385)  
  
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 956 CCCTGTCACCAACCAAGAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAAGTGCAGC 1015  
|||  
  
Qy 21 LysProCysAla 24  
|||  
Db 1016 AAGCCCTGTGCC 1027  
|||  
  
RESULT 28  
ADZ47806  
ID ADZ47806 standard; DNA; 2437 BP.  
XX  
XX ADZ47806;  
AC  
XX 30-JUN-2005 (first entry)  
XX  
XX DNA encoding human ErbB-2 variant IV.  
XX  
XX epidermal growth factor receptor-2; ErbB-2; neu; diagnosis;  
KW prognosis; cancer; cytostatic; gene; ds.  
XX  
XX Homo sapiens.  
OS  
PH Key Location/Qualifiers  
FT CDS 151..2304  
FT /\*tag= a  
FT /product= "ErbB-2 variant IV"  
XX  
XX WO200503133-A2.  
XX  
XX 14-APR-2005.  
XX  
XX 04-OCT-2004; 2004WO-US030903.  
XX  
XX 03-OCT-2003; 2003US-0507953P.  
PR 10-NOV-2003; 2003US-0518321P.  
PR 09-AUG-2004; 2004US-0599583P.  
PR 07-SEP-2004; 2004US-0607326P.  
XX

PA (COMP-) COMPUGEN LTD.  
PA (COMP-) COMPUGEN INC.  
XX  
XX Shemesh R, Oren A, Rotman G, Sela-Tavor O, Walach S;  
PI Sameach-Greenwald S, Beiman M, Eshel D, Savitsky K;  
XX  
XX WPI; 2005-285403/29.  
DR P-PSDB; ADZ47807.  
XX  
XX New isolated polynucleotide encoding at least an active portion of an  
PT ErbB-2 polypeptide, useful for diagnosing or treating diseases with  
PT aberrant expression or activity of the ErbB-2 polypeptides, such as  
PT cancer.  
XX  
XX Claim 3; SEQ ID NO 13; 246pp; English.  
PS  
XX The invention relates to the isolation of polynucleotide sequences  
CC encoding novel epidermal growth factor receptor-2 (ErbB-2, also known as  
CC HER-2 or neu) polypeptides. Also described are methods, compositions, and  
CC kits using the novel ErbB-2 sequences for the diagnosis, prognosis, and  
CC treatment of ErbB-2 related cancer. The sequences, methods, and  
CC compositions of the present invention are useful for diagnosing,  
CC prognosing, and/or treating diseases or conditions associated with  
CC aberrant expression or activity of the ErbB-2 polypeptides, such as  
CC cancer. This sequence encodes human ErbB-2 variant IV.  
XX  
SQ Sequence 2437 BP; 470 A; 806 C; 702 G; 459 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 2,53e-10 Length: 2437  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 14 Gaps: 0  
  
US-09-632-036F-6 (1-24) x ADZ47806 (1-2437)  
  
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 1096 CCCTGTCACCAACCAAGAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAAGTGCAGC 1155  
|||  
  
Qy 21 LysProCysAla 24  
|||  
Db 1156 AAGCCCTGTGCC 1167  
|||  
  
RESULT 29  
ADT50728  
ID ADT50728 standard; DNA; 2802 BP.  
XX  
XX ADT50728;  
AC  
XX 13-JAN-2005 (first entry)  
XX  
XX Cancer related nucleic acid sequence #32.  
DE  
XX ds; gene; cytostatic; Gene therapy; vaccine; diagnosis; breast; colon;  
KW lung; ovarian; prostate; cancer.  
XX  
XX Homo sapiens.  
OS  
XX WO2004092338-A2.  
PN  
XX 28-OCT-2004.  
PD  
XX 12-APR-2004; 2004WO-US011104.  
PF  
XX 11-APR-2003; 2003US-0462399P.  
PR 01-JUL-2003; 2003US-0484333P.  
XX  
XX (DIAD-) DIADEXUS INC.  
XX  
XX Macina RA, Turner LR, Sun Y, Tam A;  
PI

```
XX WPI; 2004-766851/75.
DR
XX New cancer specific nucleic acid (CaSNA) molecules, useful for
PT diagnosing, monitoring the presence of, or treating a patient with
PT breast, colon, lung, ovarian, or prostate cancer.
XX
XX Claim 1; SEQ ID NO 32; 891pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I)
CC selectively hybridizing to, or comprising at least 95% sequence identity
CC to, any of the 362 nucleotide sequences fully defined in the
CC specification. The nucleic acid molecules and polypeptides are useful for
CC diagnosing, monitoring the presence of, or treating a patient with
CC breast, colon, lung, ovarian, or prostate cancer. This sequence
CC corresponds to a nucleic acid of the invention.
XX
XX Sequence 2802 BP; 567 A; 859 C; 843 G; 532 T; 0 U; 1 Other;
SQ
Alignment Scores:
Pred. No.: 3e-10 Length: 2802
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 13 Gaps: 0
US-09-632-036F-6 (1-24) x ADT50728 (1-2802)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCCTGCACACCAAGAGGTGCACAGCAGAGGATGCACACACAGCGTGTGAGAAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167
RESULT 30
AAZ50586
ID AAZ50586 standard; DNA; 2871 BP.
XX
AC AAZ50586;
XX
DT 23-MAY-2000 (first entry)
XX
DE DC8scFv-erbB2EC fusion construct: containing tetramerisation domain.
XX
KW DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;
KW human; fusion construct; tetramerisation domain; constant domain;
KW heteroinibody; multifunctional compound; melanoma; sarcoma;
KW immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;
KW antiproliferative; prevention; treatment; malignant; haematopoietic cell;
KW lymphoma; leukaemia; solid tumour; carcinoma; ds.
XX
OS Unidentified.
OS Homo sapiens.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 10..2865
FT /*tag= a
FT /product= "DC8scFv-erbB2EC fusion protein"
FT sig_peptide 10..66
FT /*tag= b
FT /label= Leader_sequence
FT misc_feature 67..390
FT /*tag= c
FT /label= DC8scFv_light_chain_variable_region
FT misc_feature 391..435
FT /*tag= d
FT /label= Glycine-Serine-linker_DNA
FT misc_feature 436..771
FT /*tag= e
```

```
FT misc_feature /label= DC8scFv_heavy_chain_variable_region
FT 775..807
FT /*tag= f
FT /note= "5' end of human Igc3 upper hinge region with 3
FT additional nucleotides"
FT 808..924
FT /*tag= g
FT /label= Human_p53_tetramerisation_domain
FT 925..945
FT /*tag= h
FT /label= Short_peptide_linker
FT 946..2844
FT /*tag= i
FT /label= erbB2EC_domain
FT 2845..2862
FT /*tag= j
FT /label= His_tag
XX
XX WO200006605-A2.
XX
XX 10-FEB-2000.
XX
XX 28-JUL-1999; 99WO-EP005416.
XX
XX 28-JUL-1998; 98EP-00114082.
XX
XX (MICR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.
XX
XX Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;
XX WPI; 2000-195265/17.
DR P-PSDB; AAY44993.
XX
XX New multifunctional compounds useful for preventing and/or treating
PT malignant cell growth and for detection and diagnosis.
XX
XX Example 9; Fig 49; 166pp; English.
XX
XX The patent discloses heteroinibodies which are multifunctional compounds
CC producible in a mammalian host cell as a secretable and fully functional
CC heterodimer of two polypeptide chains, where one of the polypeptide
CC chains comprises, a CH1-domain (constant domain of an immunoglobulin
CC heavy chain) and the other chain comprises CL-domain (constant domain of
CC an immunoglobulin light chain). The polypeptide chains further comprise,
CC fused to the constant domains at least two (poly)peptides having
CC different receptor or ligand functions, where further at least two of the
CC different (poly)peptides lack an intrinsic affinity for one another and
CC are linked via the constant domains. The heteroinibodies have
CC cytostatic, immunostimulatory, antileukaemia and antiproliferative
CC activities. These compounds can be used for diagnosing, preventing and
CC treating malignant cell growth related to malignancies of haematopoietic
CC cells e.g. lymphomas and leukaemias, or to solid tumours e.g. carcinomas,
CC melanomas and sarcomas. The present sequence is a fusion construct
CC comprising DC8 single-chain Fv (scFv) fragment at the N-terminus,
CC extracellular domain of human erbB2 at the C-terminus and a
CC tetramerisation domain between them. This construct was prepared to find
CC out whether an oligomerisation domain characterised in bacterial
CC expression system is applicable for expression of fully functional and
CC secretable recombinant protein in mammalian host cells. This tetrameric
CC construct was not expressed as secretable and fully functional protein in
CC mammalian cells. Hence general applicability of the tetramerisation
CC domain for oligomerisation strategies in mammalian cells was ruled out
XX
XX Sequence 2871 BP; 598 A; 868 C; 834 G; 571 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3.09e-10 Length: 2871
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0
```

```
US-09-632-036F-6 (1-24) x AAZ50586 (1-2871)
QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
      1828 CCCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAAGTCGAGC 1887
QY      21 LysProCysAla 24
      1888 AAGCCCTGTGCC 1899
Db      1888 AAGCCCTGTGCC 1899
RESULT 31
ADM29351
ID      ADM29351 standard; DNA; 3188 BP.
AC      ADM29351;
XX
DT      20-MAY-2004 (first entry)
XX
DE      Human novel protein NOV20a coding sequence.
XX
KW      human; novel protein; NOV; cancer; immune associated disorder; gene; ds.
XX
OS      Homo sapiens.
XX
PN      WO2003064628-A2.
XX
PD      07-AUG-2003.
XX
PF      03-FEB-2003; 2003WO-US003401.
XX
PR      01-FEB-2002; 2002US-0353287P.
PR      01-FEB-2002; 2002US-0353301P.
PR      12-FEB-2002; 2002US-0356371P.
PR      12-FEB-2002; 2002US-0356424P.
PR      13-FEB-2002; 2002US-0356531P.
PR      20-FEB-2002; 2002US-0358239P.
PR      26-FEB-2002; 2002US-0359603P.
PR      27-FEB-2002; 2002US-0359848P.
PR      27-FEB-2002; 2002US-0359860P.
PR      15-MAR-2002; 2002US-0365049P.
PR      22-MAR-2002; 2002US-0366802P.
PR      17-MAY-2002; 2002US-0381666P.
PR      18-JUN-2002; 2002US-0389531P.
PR      19-JUN-2002; 2002US-0389910P.
PR      25-JUN-2002; 2002US-0391516P.
PR      02-JUL-2002; 2002US-0393265P.
PR      07-AUG-2002; 2002US-0401825P.
PR      09-AUG-2002; 2002US-0402395P.
PR      12-AUG-2002; 2002US-0402867P.
PR      23-AUG-2002; 2002US-0405401P.
PR      23 AUG 2002; 2002US 0405820P.
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Alsobrook JP, Bader JS, Berghs C, Burgess CE, Casman SJ;
PI      Catterton E, Chaudhuri A, Edinger SR, Ellerman K, Gerlach VL;
PI      Gorman L, Guo X, Herrmann JL, Ji W, Khrantsov NV, Li L, Miller CE;
PI      Ort T, Patturajan M, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI      Spytek KA, Vernet CAM, Zhong H, Zhong M;
XX
WPI; 2003-646149/61.
DR      P-PSDB; ADM29352.
XX
PT      New NOVX polypeptide, useful for the manufacture of a medicament for
PT      treating e.g., cancer or immune associated disorders.
XX
PS      Claim 20; SEQ ID NO 95; 606pp; English.
XX
CC      The invention comprises the amino acid and coding sequences of novel
CC      human proteins (NOV proteins). The DNA and protein sequences of the
CC      invention are useful for the manufacture of a medicament for treating a
CC      syndrome associated with a human disease comprising a pathology
CC      associated with the protein, such as: cancer or immune associated
```

```
CC disorders. The present DNA sequence encodes a NOV protein of the
CC invention.
XX
SQ      Sequence 3188 BP; 648 A; 959 C; 956 G; 625 T; 0 U; 0 Other;
      Alignment Scores:
      Pred. No.:      3.52e-10      Length:      3188
      Score:      128.00      Matches:      23
      Percent Similarity:      95.83%      Conservative:      0
      Best Local Similarity:      95.83%      Mismatches:      1
      Query Match:      96.97%      Indels:      0
      DB:      11      Gaps:      0
US-09-632-036F-6 (1-24) x ADM29351 (1-3188)
QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
      964 CCCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAAGTCGAGC 1023
Db      964 CCCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAAGTCGAGC 1023
QY      21 LysProCysAla 24
      1024 AAGCCCTGTGCC 1035
Db      1024 AAGCCCTGTGCC 1035
RESULT 32
AAA89736
ID      AAA89736 standard; DNA; 3600 BP.
XX
AC      AAA89736;
XX
DT      12-JAN-2001 (first entry)
XX
DE      Human HER-2/neu coding sequence.
XX
KW      Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW      breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW      colon cancer; ds.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..3600
FT      FT      /*tag= a
FT      FT      /product= "HER-2/neu protein"
XX
PN      WO200044899-A1.
XX
PD      03-AUG-2000.
XX
PF      28-JAN-2000; 2000WO-US002164.
XX
PR      29-JAN-1999; 99US-0117976P.
XX
PA      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM.
XX
PI      Cheever MA, Gheysen D;
XX
WPI; 2000-505976/45.
DR      P-PSDB; AAB21198, AAB21208.
XX
PT      HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT      useful for vaccinating against breast, ovarian, colon, lung and prostate
PT      cancers.
XX
PS      Disclosure; Fig 15; 128pp; English.
XX
CC      The present sequence encodes the human HER-2/neu protein. HER-2/neu is a
CC      member of the tyrosine kinase family of receptor-like glycoproteins and
CC      shows homology to the epidermal growth factor receptor (EGFR). It
CC      probably plays a part in cell growth and/or differentiation. The HER-
CC      2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a HER-
CC      2/neu extracellular domain fused to a HER-2/neu phosphorylation domain
CC      may be used to treat or prevent cancer by eliciting or enhancing an
```



CC immune response to the HER-2/neu protein. It may be used to treat  
CC malignancies such as breast, ovarian, colon, lung and prostate cancers,  
CC and may be used as an antigen to vaccinate against these neoplasias  
XX  
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.08e-10 Length: 3600  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0  
US-09-632-036F-6 (1-24) x AAA89736 (1-3600)  
QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
DB 946 C C C C T G C A C A C C A A G G T G A C A G C A G G A T G G A A C A C A G C G G T G T G A G A A G T G C A G C 1005  
QY 21 LysProCysAla 24  
DB 1006 A A G C C C T G T G C C 1017  
RESULT 33  
ID ABK86207 standard; cDNA; 3678 BP.  
XX  
AC ABK86207;  
XX  
DT 24-SEP-2002 (first entry)  
XX  
DE cDNA encoding human breast cancer antigen, Her2 variant.  
XX  
KW Human; Her2; cytostatic; antiviral; immunostimulant;  
KW cell-mediated immune response; tumour; breast cancer; virus infection;  
KW prostate cancer; colorectal cancer; pancreatic cancer; lymphoma;  
KW leukaemia; hepatitis virus; lentivirus; herpesvirus;  
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 7..3678  
FT /\*tag= a  
FT /product= "Breast cancer antigen Her2 variant"  
XX  
PN WO200240059-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 01-NOV-2001; 2001WO-US045626.  
XX  
PR 01-NOV-2000; 2000US-00704232.  
XX  
PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.  
PA (MINC/) MINCHEFF M S.  
PA (LOUK/) LOUKINOV D I.  
PA (ZOUR/) ZOURAK S.  
XX  
PI Mincheff MS, Loukinov DI, Zoubak S;  
XX  
DR WPI; 2002-527524/56.  
DR P-PSDB; AAU98923.  
XX  
PT Inducing a cell-mediated immune response against a target antigen,  
PT reducing undesired cells and stimulating presentation of an antigen by a  
PT cell, comprises administering a polynucleotide encoding a variant of an  
PT antigen.  
XX  
PS Disclosure; Page 128-134; 146pp; English.  
XX  
CC The invention relates to a method of inducing a cell-mediated immune

CC response against a cell comprising a target antigen (I) in a subject,  
CC treating a subject having undesired cells, for example tumour cells or  
CC virally infected cells (C), reducing the number of (C) in a subject, and  
CC stimulating presentation of (I) by a cell. This is done by administering  
CC a polynucleotide (II) encoding a variant of (I), so that (II) expressed  
CC in a cell and cell-mediated immune response is induced. The method can be  
CC used to treat prostate cancer, breast cancer, colorectal cancer and  
CC pancreatic cancer, as well as lymphomas and leukaemias. The method is  
CC also useful in treating chronic viral infections such as those caused by  
CC hepatitis viruses, lentiviruses (including human immunodeficiency virus  
CC (HIV)), herpesviruses and the flaviviruses and pestiviruses. The present  
CC sequence represents the coding sequence of human breast cancer antigen,  
CC Her2 variant, used as a target antigen in the method of the invention  
XX  
SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.19e-10 Length: 3678  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-6 (1-24) x ABK86207 (1-3678)  
QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
DB 856 C C C C T G C A C A C C A A G G T G A C A G C A G G A T G G A A C A C A G C G G T G T G A G A A G T G C A G C 915  
QY 21 LysProCysAla 24  
DB 916 A A G C C C T G T G C C 927  
RESULT 34  
ID ADB67620 standard; DNA; 3765 BP.  
XX  
AC ADB67620;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human epidermal growth factor receptor 2 coding sequence.  
XX  
KW ds; gene; cytostatic; human epidermal growth factor receptor-3; HER-3;  
KW heregulin; HER2; tyrosine kinase activity; cancer.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3765  
FT /\*tag= a  
FT /product= "Human epidermal growth factor receptor 2"  
FT /note= "no stop codon given at the 3' end"  
XX  
PN WO2003011897-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 29-JUL-2002; 2002WO-US023963.  
XX  
PR 27-JUL-2001; 2001US-0308341P.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Singer E, Landgraf R, Slamon DJ, Eisenberg D;  
XX  
DR WPI; 2003-300482/29.  
DR P-PSDB; ADB67621.  
XX  
PT Novel human epidermal growth factor receptor 3 variant as agonist or  
PT antagonist of HER3 receptor, for diagnosis/treatment of cells or  
PT pathological conditions associated with aberrant expression of heregulin

PT or HER3.  
 XX Disclosure; Page 80-81; 137pp; English.  
 PS  
 XX The invention relates to a non-naturally occurring human epidermal growth  
 CC factor receptor (HER)-3 variant polypeptide comprising amino acids 19-329  
 CC or 20-329 of the 1342 amino acid HER3 polypeptide (ADB67617) or a  
 CC sequence which differs from native HER3 polypeptide and having amino acid  
 CC substitutions at residues E43, N44, K51, E64, V66 and V110 of S1, is new.  
 CC The variant HER-3 specifically binds to the heregulin polypeptide  
 CC (ADB67619), exhibits an impaired ability to interact with HER2  
 CC polypeptide (ADB67621), or has an ability to inhibit the interaction  
 CC between wild-type HER3 and heregulin. The polypeptide is useful for  
 CC identifying a compound which specifically binds to heregulin binding  
 CC domain in a HER3 variant polypeptide. The method further involves  
 CC determining whether the test compound inhibits or enhances the heregulin  
 CC induced tyrosine kinase activity associated with a HER3 polypeptide. The  
 CC polypeptide is also useful for determining whether a test compound  
 CC modulates the interaction between a heregulin polypeptide, and the  
 CC variant HER-3 polypeptide. The HER-3 polypeptide is also useful for  
 CC inhibiting the interaction between a heregulin polypeptide and HER3  
 CC polypeptide, e.g. for treating cancer. The polypeptide is also useful for  
 CC stimulating or activating HER3 receptor. This sequence represents the  
 CC coding sequence for the wild type human HER-2 polypeptide.  
 XX  
 SQ Sequence 3765 BP; 757 A; 1170 C; 1120 G; 718 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 4,31e-10 Length: 3765  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-632-036F-6 (1-24) x ADB67620 (1-3765)  
 QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 Db 946 CCCTTGCACACCAACAGAGGTGACAGCAGAGGATGGACACACAGCGGTGTGAGAGTGCAGC 1005  
 QY 21 LysProCysAla 24  
 Db 1006 AAGCCCTGTGCC 1017  
 RESULT 35  
 ADZ72379  
 ID ADZ72379 standard; cDNA; 3765 BP.  
 XX  
 AC ADZ72379;  
 XX  
 DT 14-JUL-2005 (first entry)  
 XX  
 DE DNA encoding human epidermal growth factor receptor-2, HER2.  
 XX  
 KW Aptamer; epidermal growth factor receptor-2; HER2; tumor; drug screening;  
 KW cancer; cytostatic; diagnosis; carcinoma; lymphoma; sarcoma; leukemia;  
 KW ss; gene.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CD 1..3765  
 FT /\*tag=a  
 FT /product="HER2"  
 XX  
 PN WO2005040339-A2.  
 XX  
 XX 06-MAY-2005.  
 PD  
 XX 16-JUL-2004; 2004WO-US023039.  
 PF  
 XX 18-JUL-2003; 2003US-0488679P.  
 PR

XX (REGC ) UNIV CALIFORNIA.  
 PA  
 XX Chen CB, Landgraf R;  
 PI  
 XX WPI; 2005-333498/34.  
 DR P-P8DB; ADZ72380.  
 XX  
 PT New aptamers to human epidermal growth factor receptor-3, useful for  
 PT diagnosing, treating or preventing cancer, e.g. carcinoma, lymphoma,  
 PT blastoma, sarcoma, and leukemia.  
 PT  
 PS Disclosure; SEQ ID NO 5; 74pp; English.  
 XX  
 XX The present invention relates to compositions of nucleic acid aptamers  
 CC that bind to human epidermal growth factor receptor-3 (HER3) comprising  
 CC the protein sequence given as SEQ ID NO: 2 (see ADZ72376). HER3 is a  
 CC receptor tyrosine kinase which are involved in a broad spectrum of cell  
 CC growth and differentiation events. Overexpression of members of this  
 CC receptor family such as epithelial growth factor receptor and, HER3 and  
 CC HER2, is associated with a variety of solid tumor malignancies. SELEX  
 CC (Systematic Evolution of Ligands by Exponential enrichment), was used to  
 CC select RNA aptamers against the oligomeric states of the extracellular  
 CC domains (ECD) of HER3. A number of specific RNA aptamers and methods of  
 CC making these aptamers are disclosed. One of the aptamers, A30 binds with  
 CC high affinity to a limited number of binding sites in the oligomeric  
 CC state of HER3ECD. High affinity binding of A30 inhibits heregulin-  
 CC dependent tyrosine phosphorylation of HER2 as well as heregulin induced  
 CC growth of MCF7 cells. This aptamer can be useful as a tool for analysis  
 CC of receptor interactions and may serve as a lead compound for the  
 CC development of inhibitors against overexpressed receptor tyrosine kinases  
 CC in pathologies associated with HER3 overexpression, such as cancer.  
 CC Disclosed is one such aptamer given as SEQ ID NO: 19 (see ADZ72393),  
 CC others comprise SEQ ID NOS: 7, 12, 15-18. The aptamers further comprise a  
 CC fluorine moiety or an amino moiety. They form a hairpin loop structure  
 CC and further comprise a stem structure comprised of at least 1-6 base  
 CC pairs. Also disclosed a nucleic acid (SEQ ID NO:1) encoding a HER3  
 CC polypeptide. The nucleic acid, a composition, and an analysis method are  
 CC useful for diagnosing, treating or preventing cancer, e.g. carcinoma,  
 CC lymphoma, blastoma, sarcoma, and leukemia. The present sequence is a DNA  
 CC sequence encoding HER2.  
 XX  
 SQ Sequence 3765 BP; 757 A; 1170 C; 1120 G; 718 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 4,31e-10 Length: 3765  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-632-036F-6 (1-24) x ADZ72379 (1-3765)  
 QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 Db 946 CCCTTGCACACCAACAGAGGTGACAGCAGAGGATGGACACACAGCGGTGTGAGAGTGCAGC 1005  
 QY 21 LysProCysAla 24  
 Db 1006 AAGCCCTGTGCC 1017  
 RESULT 36  
 AAT40739  
 ID AAT40739 standard; cDNA; 3768 BP.  
 XX  
 AC AAT40739;  
 XX  
 DT 01-JAN-1997 (first entry)  
 XX  
 DE HER-2/neu oncogene.  
 XX  
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;

KW breast cancer; ovary cancer; colon cancer; lung cancer; prostate cancer;  
KW genetic immunisation; tumour; vaccine; vector; ss.  
XX Homo sapiens.

XX  
FH Key Location/Qualifiers  
PH CDS 1..3765  
FT /\*tag= b  
FT /note= "nucleotides 2026-3765 (claim 1) code for HER-  
FT 2/neu intracellular domain"  
XX

XX W09J0514 AL.

XX 03-OCT-1996.

XX 28-MAR-1996; 96WO-US001689.

XX 31-MAR-1995; 95US-00414417.

XX (UNIW ) UNIV WASHINGTON.

XX Cheever MA, Disis ML;

XX WPI; 1996-455361/45.

XX P-PSDB; AAW01111.

XX DNA encoding HER-2/neu poly:peptide(s) - used for prevention or treatment  
XX of malignancies with which the HER-2/neu oncogene is associated.

XX Claim 1; Page 49-56; 71pp; English.

XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or c-  
XX erbB2) protein (AAW01111). The oncogene is overexpressed in various  
XX cancers, including breast, ovarian, colon, lung and prostate, and appears  
XX to induce malignancies through quantitative mechanisms that result from  
XX increased or deregulated expression of an essentially normal gene  
XX product. Nucleotides 2026-3765 of the cDNA sequence code for the  
XX intracellular domain (lys676-val1255) of the HER-2/neu protein, which is  
XX useful for immunisation against malignancy. Nucleic acids can be used to  
XX direct expression of the intracellular domain in transformed host cells,  
XX or are used, alone or in a viral vector, for genetic immunisation of an  
XX animal

XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.32e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x AAT40739 (1-3768)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005

Qy 21 LysProCysAla 24

Db 1006 AAGCCCTGTGCC 1017

RESULT 37

AAAX01912

ID AAX01912 standard; DNA; 3768 BP.

XX AAX01912;

XX 21-APR-1999 (first entry)

XX Human HER-2/neu oncogene DNA.

XX

KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;  
KW malignancy; treatment; tumour; ss.  
XX Homo sapiens.

XX  
FH Key Location/Qualifiers  
PH CDS 1..3768  
FT /\*tag= a  
FT /product= "HER-2/neu"  
FT /note= "oncogene"  
FT 2026..3765  
FT misc\_feature  
FT /\*tag= b  
FT /note= "region which elicits immune response"  
XX

XX US5869445-A.

XX 09-FEB-1999.

XX 01-APR-1996; 96US-00625101.

XX 17-MAR-1993; 93US-00033644.

XX 12-AUG-1993; 93US-00106112.

XX 31-MAR-1995; 95US-00414417.

XX (UNIW ) UNIV WASHINGTON.

XX Cheever MA, Disis ML;

XX WPI; 1999-152835/13.

XX P-PSDB; AAW92406.

XX Use of HER-2/neu polypeptides - for eliciting an immune response to an  
XX HER-2/neu associated malignancy, particularly for treating or preventing  
XX tumours.

XX Claim 1a; Col 23-32; 26pp; English.

XX This sequence encodes the human HER-2/neu oncogene protein. A fragment of  
XX this protein is used in a method for eliciting or enhancing an immune  
XX response to HER-2/neu protein. The polypeptide can stimulate T cells and  
XX B cells to produce an immune response to the HER-2/neu protein. The  
XX method can be used for immunisation against a malignancy in which the HER  
XX -2/neu oncogene is associated and in the treatment of an existing tumour,  
XX or to prevent tumour occurrence or reoccurrence

XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.32e-10 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x AAX01912 (1-3768)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005

Qy 21 LysProCysAla 24

Db 1006 AAGCCCTGTGCC 1017

RESULT 38

AAA09455

ID AAA09455 standard; DNA; 3768 BP.

XX AAA09455;

XX 10-AUG-2000 (first entry)

XX

XX





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 22:40:44 ; Search time 93.8182 Seconds  
(without alignments)  
454.725 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLNQEVTAEDGTQRAEKSCPKCA 24

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

-Q=/cgn2\_1/USPIO\_spoel/US09632036/runat 02122005 104141 25840/app query.fasta 1.398

-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=100 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USPR=US09632036 @CGN 1 1 188 @runat 02122005 104141 25840 -NCPUs=6 -ICPU=3

-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/PGTUS COMB.seq.\*

7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*

8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*

9: /cgn2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	97.0	1872	3	US-08-422-108-2
2	128	97.0	1872	3	US-08-422-734-2
3	128	97.0	2385	2	US-09-146-283-3
4	128	97.0	2385	3	US-08-579-823A-3
5	128	97.0	2385	3	US-09-344-195-3
6	128	97.0	3768	2	US-08-625-101-1
7	128	97.0	3768	2	US-08-356-786-1
8	128	97.0	3768	3	US-09-811-115-2
9	128	97.0	3768	3	US-09-167-516-1
10	128	97.0	4473	2	US-09-048-804-1
11	128	97.0	4473	3	US-09-056-105-26
12	128	97.0	4473	3	US-09-663-83A-3
13	128	97.0	4473	3	US-09-441-411-5
14	128	97.0	4530	2	US-08-229-515A-9
15	128	97.0	4530	3	US-08-645-865-9
16	128	97.0	4530	3	US-09-167-322-4
17	128	97.0	4530	3	US-09-527-487-1
18	128	97.0	4530	3	US-09-877-177A-11
19	128	97.0	9274	3	US-09-811-115-1
20	114	86.4	3955	2	US-08-229-515A-14
21	114	86.4	3955	2	US-08-645-865-14
22	54	40.9	601	3	US-09-949-016-195823
23	54	40.9	251672	3	US-09-949-016-17296
24	54	40.9	251682	3	US-09-949-016-11973
25	53.5	40.5	1288	3	US-09-270-767-232
26	53.5	40.5	1288	3	US-09-270-767-15514
27	52	39.4	471	3	US-09-235-451-10
28	52	39.4	471	3	US-09-978-303-10
29	52	39.4	1062	3	US-09-902-540-6754
30	52	39.4	1318	2	US-08-809-494A-5
31	52	39.4	1318	3	US-09-352-302-5
32	52	39.4	2637	3	US-09-949-016-5623
33	52	39.4	3707	3	US-09-902-540-549
34	52	39.4	6975	3	US-09-902-540-2386
35	52	39.4	17315	3	US-09-902-540-1103
36	52	39.4	27227	3	US-09-949-016-17365
37	52	39.4	141248	3	US-09-949-016-12241
38	52	39.4	143248	3	US-09-949-016-16652
39	51	38.6	1763	3	US-09-385-219A-13
40	50	37.9	601	3	US-09-949-016-137588
41	50	37.9	601	3	US-09-949-002-3430
42	50	37.9	601	3	US-09-949-002-9447
43	50	37.9	2236	3	US-09-897-367-1
44	50	37.9	2236	3	US-09-918-909A-1
45	50	37.9	2351	3	US-09-016-434-1337
46	50	37.9	26664	3	US-09-564-805-28
47	50	37.9	26664	3	US-09-434-382-28
48	50	37.9	29558	3	US-09-949-016-15607
49	50	37.9	51339	3	US-09-949-002-677
50	50	37.9	51340	3	US-09-949-002-819
51	50	37.9	139150	3	US-09-949-016-17398
52	50	37.9	139577	3	US-09-949-016-12879
53	49.5	37.5	76164	3	US-09-949-016-12288
54	49.5	37.5	76165	3	US-09-949-016-14005
55	49	37.1	458	3	US-09-270-767-943
56	49	37.1	458	3	US-09-270-767-16225
57	49	37.1	532	3	US-09-902-540-1555
58	49	37.1	1593	3	US-09-676-610B-25
59	49	37.1	1868	2	US-08-658-883B-1
60	49	37.1	1868	3	US-09-676-610B-26
61	49	37.1	2801	3	US-09-291-922-7
62	49	37.1	3633	2	US-09-715-249-1
63	49	37.1	5532	2	US-08-475-035-3
64	49	37.1	5532	3	US-09-676-610B-17
65	49	37.1	169998	3	US-09-676-610B-24
66	49	37.1	17496	3	US-09-877-177A-10
67	48.5	36.7	601	3	US-09-949-016-128332
68	48.5	36.7	601	3	US-09-949-016-128333
69	48.5	36.7	601	3	US-09-949-016-128334
70	48.5	36.7	601	3	US-09-949-016-128335
71	48.5	36.7	601	3	US-09-949-016-128336
72	48.5	36.7	961	3	US-09-949-016-128336
73	48.5	36.7	4202	3	US-09-511-625B-17
74	48.5	36.7	4202	3	US-09-904-615-56
75	48.5	36.7	4202	3	US-10-054-988-56
76	48.5	36.7	83516	3	US-09-511-625B-5
77	48	36.4	601	3	US-09-949-016-15378
78	48	36.4	601	3	US-09-949-016-65682
79	48	36.4	601	3	US-09-949-016-65683
80	48	36.4	601	3	US-09-949-016-177055
81	48	36.4	601	3	US-09-949-016-177056
82	48	36.4	601	3	US-09-949-002-3429
83	48	36.4	601	3	US-09-949-002-9446

```
c 83 48 36.4 653 3 US-09-902-540-1382
c 84 48 36.4 783 3 US-09-252-991A-7428
c 85 48 36.4 1483 3 US-09-262-749-1
c 86 48 36.4 1557 3 US-09-719-978-1
c 87 48 36.4 1958 3 US-09-570-454-1
c 88 48 36.4 1958 3 US-09-867-521-1
c 89 48 36.4 2421 3 US-09-902-540-7795
c 90 48 36.4 2541 3 US-09-252-991A-7209
c 91 48 36.4 3099 3 US-09-252-991A-7487
c 92 48 36.4 5717 3 US-09-902-540-776
c 93 48 36.4 6737 3 US-09-453-702B-76
c 94 48 36.4 6737 3 US-10-114-170-76
c 95 48 36.4 6775 3 US-09-620-312D-289
c 96 48 36.4 7280 3 US-09-902-540-827
c 97 48 36.4 9862 3 US-09-691-861A-3
c 98 48 36.4 10320 3 US-09-949-016-12117
c 99 48 36.4 10321 3 US-09-949-016-16155
c 100 48 36.4 19237 3 US-09-949-016-13666

ALIGNMENTS

RESULT 1
US-08-422-108-2
; Sequence 2, Application US/08422108
; Patent No. 6015567
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,108
; FILING DATE: 14-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-422-108-2

ALIGNMENTS

RESULT 2
US-08-422-734-2
; Sequence 2, Application US/08422734
; Patent No. 633169
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,734
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422108
; FILING DATE: 14-Apr-1995
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-422-734-2

Alignment Scores: 4.43e-10 Length: 1872
Pred. No.:
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-422-108-2 (1-1872)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrClnArgAlaGluLysCysSer 20
Db 883 CCCUGCACAACCAAGAGUGAGCAGCAGAGGAUGGAACACAGCGUGUGAGAGAGCAGC 942
QY 21 LysProCysAla 24
Db 943 AAGCCCCUGGCC 954

Alignment Scores: 4.43e-10 Length: 1872
Pred. No.:
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0
```



Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	3	Gaps:	0

US-09-632-036F-6 (1-24) x US-08-422-734-2 (1-1872)

Qy	1	ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer	20
Db	883	CCCCUGACACCAAGAGUGACACGAGGAUGAACACACGCGGUGAGAAGUGCAGC	942
Qy	21	LysProCysAla	24
Db	943	AAGCCCCUGGCC	954

### RESULT 3

US-09-146-283-3  
 ; Sequence 3, Application US/09146283  
 ; Patent No. 5976546  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Ruegg, Curtis L.  
 ; APPLICANT: Wu, Hongyu  
 ; TITLE OF INVENTION: Immunostimulatory Compositions  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: 350 Cambridge Ave. Suite 250  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA

	Alignment Scores:		
Pred. No.:	Score:	Length:	2385
	5.87e-10	Matches:	23
	128.00	Conservative:	0
Percent Similarity:	95.83%	Mismatches:	1
Best Local Similarity:	95.83%	Indels:	0
Query Match:	96.97%	Gaps:	0
DB:	2		

US-09-632-036F-6 (1-24) x US-09-146-283-3 (1-2385)

Qy	1	ProLeuHiAAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer	20
Dd	956	CCCTGTCAACCAAGAGGTGCAGCAGAGGATGGAAACACAGCGTGTGAGAAGTCAGC	1015
Qy	21	LysProCysAla	24
Dd	1016	AAGCCCTGTGCC	1027

## RESULT 4

```

RES001.4
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruesg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3

```

Alignment Scores:	5.87e-10	Length:	2385
Pred. No.:	128.00	Matches:	23
Score:	95.83%	Conservative:	0
Percent Similarity:	95.83%	Mismatches:	1
Best Local Similarity:	96.97%	Indels:	0
Query Match:	3	Gaps:	0
DB:			

US-09-632-036F-6 (1-24) x US-08-579-823A-3 (1-2385)

Qy	1	ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer	20
Db	956	CCCCCGCACCAACAGAGGTGACAGAGGATGGAAACACACGCGGTGTGAGAAGTGCACG	1015
Qy	21	LysProCysAla	24
Db	1016	AAGCCCTGTGCC	1027

## RESULT 5

```
US 09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laue, Reiner
;             Ruegg, Curtis L.
;             Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSP-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
Pred. No.: 5.87e-10 Length: 2385
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-344-195-3 (1-2385)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 956 CCCCTGCACAAACCAAGAGGTGACAGCAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1015
QY 21 LysProCysAla 24
Db 1016 AAGCCCTGTGCC 1027

RESULT 6
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
```

## GENERAL INFORMATION:

```
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 1e-09 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-625-101-1 (1-3768)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1005
QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 7
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786

FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3768 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3768

OTHER INFORMATION: /note= "product = "cerB-b2"  
US-08-356-786-1

Alignment Scores:  
Pred. No.: 1e-09 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-356-786-1 (1-3768)

Oy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGAGTGGAAACACAGCGGTGTGAGAGTGCAGC 1005

Oy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

RESULT 8

US-09-811-115-2  
Sequence 2, Application US/09811115  
Patent No. 6632979  
GENERAL INFORMATION:  
APPLICANT: Erickson, Sharon  
APPLICANT: Schwall, Ralph  
APPLICANT: King, Kathleen  
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
FILE REFERENCE: GENENT.034A  
CURRENT APPLICATION NUMBER: US/09/811,115  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/189,844  
PRIOR FILING DATE: 2000-03-16  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Fast-SEQ for Windows Version 4.0  
SEQ ID NO 2

LENGTH: 3768  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-811-115-2

Alignment Scores:  
Pred. No.: 1e-09 Length: 3768  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-811-115-2 (1-3768)

Oy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGAGTGGAAACACAGCGGTGTGAGAGTGCAGC 1005

Oy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

RESULT 9

US-09-167-516-1

Sequence 1, Application US/09167516  
Patent No. 6953573

GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/167,516

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/625,101

FILING DATE: 01-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C7

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3768 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3765  
US-09-167-516-1

Alignment Scores:

```
Pred. No.: 1e-09 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-167-516-1 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 946 CCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1005
|||||

QY 21 LysProCysAla 24
|||||
Db 1006 AAGCCCTGTGCC 1017
|||||

RESULT 10
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
US-09-048-804-1

Alignment Scores:
Pred. No.: 1.22e-09 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-048-804-1 (1-4473)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 1120 CCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1179
|||||

QY 21 LysProCysAla 24
|||||
```

```
Db 1180 AAGCCCTGTGCC 1191
|||||

RESULT 11
US-09-056-105-26
; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; TITLE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-26

Alignment Scores:
Pred. No.: 1.22e-09 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-056-105-26 (1-4473)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||||
Db 1120 CCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1179
|||||

QY 21 LysProCysAla 24
|||||
Db 1180 AAGCCCTGTGCC 1191
|||||

RESULT 12
US-09-663-834A-3
; Sequence 3, Application US/09663834A
; Patent No. 6613567
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowbert
; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-2 EXPRESSION
; FILE REFERENCE: RTS-0033
; CURRENT APPLICATION NUMBER: US/09/663,834A
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175) ... (3942)
US-09-663-834A-3

Alignment Scores:
Pred. No.: 1.22e-09 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-663-834A-3 (1-4473)
```





Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-877-177A-11 (1-4530)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 1155  
Qy 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167

## RESULT 19

US-09-811-115-1  
; Sequence 1, Application US/09811115  
; Patent No. 6632979  
; GENERAL INFORMATION:  
; APPLICANT: Erickson, Sharon  
; APPLICANT: Schwall, Ralph  
; APPLICANT: King, Kathleen  
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
; FILE REFERENCE: GENENT.034A  
; CURRENT APPLICATION NUMBER: US/09/811.115  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/189,844  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 9274  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Vector Sequence  
US-09-811-115-1

Alignment Scores:  
Pred. No.: 2,85e-09 Length: 9274  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-811-115-1 (1-9274)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 2676 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGC 2735  
Qy 21 LysProCysAla 24  
Db 2736 AAGCCCTGTGCC 2747

## RESULT 20

US-08-229-515A-14  
; Sequence 14, Application US/08229515A  
; Patent No. 5518885  
; GENERAL INFORMATION:  
; APPLICANT: RAZIUDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/229,515A  
; FILING DATE: 19 APR 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-688-0770  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3955 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-229-515A-14

Alignment Scores:  
Pred. No.: 1,44e-07 Length: 3955  
Score: 114.00 Matches: 21  
Percent Similarity: 91.67% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 2  
Query Match: 86.36% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-229-515A-14 (1-3955)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 974 CCCCCGAATAACCAAGAGGTACAGCTGAGGACGGAACACACAGCGTTGTGTGAGAAATGCAGC 1033  
Qy 21 LysProCysAla 24  
Db 1034 AAGCCCTGTGCT 1045

## RESULT 21

US-08-645-865-14  
; Sequence 14, Application US/08645865  
; Patent No. 5654406  
; GENERAL INFORMATION:  
; APPLICANT: RAZIUDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/645,865  
; FILING DATE: 14 MAY 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:

```
/ TELEPHONE: 404-688-0770
/ TELEFAX: 404-688-9880
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3955 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 1.44e-07 Length: 3955
Score: 114.00 Matches: 21
Percent Similarity: 91.67% Conservative: 1
Best Local Similarity: 87.50% Mismatches: 2
Query Match: 86.36% Indels: 0
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-645-865-14 (1-3955)
>Y 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 974 CCCCCGAATTAACCAAGAGGTACAGCTGAGGACGGAACACAGCGTTGTGAGAAATCGC 1033

QY 21 LysProCysAla 24
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1034 AAGCCCTGTGCT 1045

RESULT 22
US-09-949-016-195823
/ Sequence 195823, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 195823
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-195823

Alignment Scores:
Pred. No.: 22.6 Length: 601
Score: 54.00 Matches: 11
Percent Similarity: 65.00% Conservative: 2
Best Local Similarity: 55.00% Mismatches: 7
Query Match: 40.91% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-949-016-195823 (1-601)
QY 5 GlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 391 CAGGAGCTCACAGCTGAGATGGGTTTCAGAAATGAAGTAAAGCCCGAGTGTGCC 450

RESULT 23
US-09-949-016-17296/c
/ Sequence 17296, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17296
/ LENGTH: 251682
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(251682)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17296

Alignment Scores:
Pred. No.: 2.56e+04 Length: 251682
Score: 54.00 Matches: 11
Percent Similarity: 65.00% Conservative: 2
Best Local Similarity: 55.00% Mismatches: 7
Query Match: 40.91% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-949-016-17296 (1-251672)
QY 5 GlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 77003 CAGGAGCTCACAGCTGAGATGGGTTTCAGAAATGAAGTAAAGCCCGAGTGTGCC 76944

RESULT 24
US-09-949-016-11973/c
/ Sequence 11973, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11973
/ LENGTH: 251682
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(251682)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11973

Alignment Scores:
Pred. No.: 2.56e+04 Length: 251682
Score: 54.00 Matches: 11
Percent Similarity: 65.00% Conservative: 2
Best Local Similarity: 55.00% Mismatches: 7
Query Match: 40.91% Indels: 0
DB: 3 Gaps: 0

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17296
/ LENGTH: 251672
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(251672)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17296

Alignment Scores:
Pred. No.: 2.56e+04 Length: 251672
Score: 54.00 Matches: 11
Percent Similarity: 65.00% Conservative: 2
Best Local Similarity: 55.00% Mismatches: 7
Query Match: 40.91% Indels: 0
DB: 3 Gaps: 0
```





```
; TYPE: DNA
; ORGANISM: M. musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(471)
; OTHER INFORMATION: n = A,T,C or G
US-09-978-303-10

Alignment Scores:
Pred. No.: 34.4 Length: 471
Score: 52.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 39.39% Indels: 2
DB: 3 Gaps: 1

US-09-632-036F-6 (1-24) x US-09-978-303-10 (1-471)

QY 8 ThrAlaGluAspGlyThrGlnArgAlaGluLys-----CysSerLysProCysala 24
Db 30 ACACGCGAGTTCAAAGACCCAGAGACGCTGTCTGCTCAAAGCCATGCTCA 86

RESULT 29
US-09-902-540-6754/c
; Sequence 6754, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6754
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6754

Alignment Scores:
Pred. No.: 88.6 Length: 1062
Score: 52.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 39.39% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-902-540-6754 (1-1062)

QY 3 HisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysPro 22
Db 1035 CACGATCAGCGGTCCACCTCCATCGGGGCGCGACGTCGGGAGGAGTGAATCATCCG 976

RESULT 30
US-08-809-494A-5
; Sequence 5, Application US/08809494A
; Patent No. 5962260
; GENERAL INFORMATION:
; APPLICANT: Sawamura, Tatsuya
; APPLICANT: Maesaki, Tomoo
; TITLE OF INVENTION: Modified Low-Density Lipoprotein
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAlulay Fisher Nissen Goldberg & Kiel
; STREET: 261 Madison Avenue
; CITY: New York
; STATE: NY
```

```
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,494A
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-321705
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-214206
; FILING DATE: 31-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldberg, Jules E
; REGISTRATION NUMBER: 24408
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 986-4090
; TELEFAX: 212 818-9479
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Lung, placenta
; IMMEDIATE SOURCE:
; LIBRARY: Human lung cDNA
; CLONE: lambdaBLOX-1
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 66..125
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 949..1309
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 127..948
; US-08-809-494A-5
Alignment Scores:
Pred. No.: 114 Length: 1318
Score: 52.00 Matches: 8
Percent Similarity: 59.09% Conservative: 5
Best Local Similarity: 36.36% Mismatches: 9
Query Match: 39.39% Indels: 0
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x US-08-809-494A-5 (1-1318)

QY 2 LeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLys 21
Db 493 CTTCAACCACCAAGATCTGAATCTCCAAGAAACACTGAAGAGAGATAGCAAAATTGTTCAAGCT 552

QY 22 ProCys 23
Db 553 CCTTGT 558

RESULT 31
US-09-352-302-5
; Sequence 5, Application US/09352302
; Patent No. 6197937
; GENERAL INFORMATION:
```

APPLICANT: Sawamura, Tateuya  
APPLICANT: Masaki, Tomoo  
TITLE OF INVENTION: Modified Low-Density Lipoprotein  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
STREET: 261 Madison Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/352.302  
FILING DATE: 12-JUL-1999

CLASSIFICATION:  
PRIOR APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E.  
REGISTRATION NUMBER: 24408

REFERENCE/DOCKET NUMBER: JG-VY-4363PCT/D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Lung, placenta  
IMMEDIATE SOURCE:  
LIBRARY: Human lung cDNA  
CLONE: lambdaHLOX-1

FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 66..125  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 949..1309  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 127..948

US-09-352-302-5

Alignment Scores:

Pred. No.: 114

Score: 52.00

Percent Similarity: 59.09%

Best Local Similarity: 36.36%

Query Match: 39.39%

DB: 3

Length: 1318

Matches: 8

Conservative: 5

Mismatches: 9

Indels: 0

Gaps: 0

US-09-632-036f-6 (1-24) x US-09-352-302-5 (1-1318)

Qy 2 LeuHieAsnGlnGluValThrAlaGluAAspGlyThrGlnArgAlaGluLysCysSerLys 21

Db 493 CTTTACCACCAAGATCTGAATCTTCCAGAAACACTGAAGAGATAGCAAAATTTGTCAGCT 552

Qy 22 ProCys 23  
Db 553 CCTTGT 558

RESULT 32

US-09-949-016-5623  
Sequence 5623, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5623  
LENGTH: 2637  
TYPE: DNA  
ORGANISM: Human

US-09-949-016-5623

Alignment Scores:

Pred. No.: 255

Score: 52.00

Percent Similarity: 66.67%

Best Local Similarity: 47.62%

Query Match: 39.39%

DB: 3

Length: 2637

Matches: 10

Conservative: 7

Mismatches: 4

Indels: 0

Gaps: 0

US-09-632-036f-6 (1-24) x US-09-949-016-5623 (1-2637)

Qy 1 ProLeuHieAsnGlnGluValThrAlaGluAAspGlyThrGlnArgAlaGluLysCysSer 20

Db 782 CCTTTGAAAGTCAGAGAGAGTGCAGAGATGTCCAGAGATCGCAGAGATGTGTGC 841

Qy 21 Lys 21

Db 842 AAA 844

RESULT 33

US-09-902-540-549/c

Sequence 549, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 549

LENGTH: 3707

TYPE: DNA

ORGANISM: Myxococcus xanthus

US-09-902-540-549

Alignment Scores:

Pred. No.: 380

Score: 52.00

Length: 3707

Matches: 9



```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12241
; LENGTH: 141248
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(141248)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12241

Alignment Scores:
Pred. No.: 2.62e+04 141248
Score: 52.00 13
Percent Similarity: 52.94% 13
Best Local Similarity: 38.24% 5
Query Match: 39.39% 6
DB: 3 10
Gaps: 1

US-09-632-036F-6 (1-24) x US-09-949-016-12241 (1-141248)

Qy 1 ProLeuHisAenGlnGluValThrAla-----Glu 10
Db 23393 CCATCCAGAAACAGAAAGTGACTGCTTATACTGGACGCTGAGCCAGTGGGCCAACCA 23334

Qy 11 AspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24
Db 23333 GACGGACCACAGAAAGCTAAGAAACCAACAGGCCAGGACCA 23292

RESULT 38
US-09-949-016-16652/c
; Sequence 16652, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16652
; LENGTH: 143248
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(143248)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16652

Alignment Scores:
Pred. No.: 2.66e+04 143248
Score: 52.00 13
Percent Similarity: 52.94% 5
Best Local Similarity: 38.24% 6
Query Match: 39.39% 10
DB: 3 1
Gaps: 1

US-09-632-036F-6 (1-24) x US-09-949-016-16652 (1-143248)

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12241
; LENGTH: 141248
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(141248)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12241

Alignment Scores:
Pred. No.: 2.62e+04 141248
Score: 52.00 13
Percent Similarity: 52.94% 5
Best Local Similarity: 38.24% 6
Query Match: 39.39% 10
DB: 3 1
Gaps: 1

US-09-632-036F-6 (1-24) x US-09-949-016-12241 (1-141248)

Qy 1 ProLeuHisAenGlnGluValThrAla-----Glu 10
Db 23393 CCATCCAGAAACAGAAAGTGACTGCTTATACTGGACGCTGAGCCAGTGGGCCAACCA 23334

Qy 11 AspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24
Db 23333 GACGGACCACAGAAAGCTAAGAAACCAACAGGCCAGGACCA 23292

RESULT 39
US-09-385-219A-13/c
; Sequence 13, Application US/09385219A
; Patent No. 6720181
; GENERAL INFORMATION:
; APPLICANT: Chiau, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219A
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-385-219A-13

Alignment Scores:
Pred. No.: 227 1763
Score: 51.00 9
Percent Similarity: 72.22% 4
Best Local Similarity: 50.00% 5
Query Match: 38.64% 0
DB: 3 0
Gaps: 0

US-09-632-036F-6 (1-24) x US-09-385-219A-13 (1-1763)

Qy 3 HisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1567 CACATCAGAAATAAACAACCTCGAGATCAGCACCCCAAGGAGTGACATCTGTAGT 1514

RESULT 40
US-09-949-016-137588
; Sequence 137588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137588
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-137588
```

Alignment Scores:  
Pred. No.: 92.1 Length: 601  
Score: 50.00 Matches: 10  
Percent Similarity: 71.43% Conservative: 0  
Best Local Similarity: 71.43% Mismatches: 4  
Query Match: 37.88% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-949-016-137588 (1-601)

Qy 9 AlaGluAspGlyThrGlnArgAlaGluLysCysSerLysPro 22  
Db 554 GCTGAGGCGAGAAACCCAGGAGCGGAAGATTGCAGTCACCCG 595

Search completed: December 11, 2005, 01:04:00  
Job time : 159.818 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 11, 2005, 00:00:23 ; Search time 470.182 Seconds  
(without alignments)  
422.103 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLHNQVTAEDGTORAEKCKPCA 24

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09632036/runat\_02122005\_104143\_25924/app\_query.fasta\_1.398  
-DB=PublishedApplications\_NA\_Main -QEMT=fastcap -SUFFIX=rnpbm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosome2  
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09632036 @C@N 1 1 1134 runat\_02122005\_104143\_25924  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOC=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	97.0	471	3	US-09-878-134-68
2	128	97.0	1260	9	US-10-775-204-1475
3	128	97.0	1320	9	US-10-956-373-25
4	128	97.0	1944	9	US-10-956-373-3
5	128	97.0	2061	9	US-10-956-373-15
6	128	97.0	2132	6	US-10-412-804A-3
7	128	97.0	2149	6	US-10-412-804A-9
8	128	97.0	2164	6	US-10-412-804A-5
9	128	97.0	2164	6	US-10-412-804A-5
10	128	97.0	2164	6	US-10-412-804A-5
11	128	97.0	2164	6	US-10-412-804A-5
12	128	97.0	2164	6	US-10-412-804A-5
13	128	97.0	2164	6	US-10-412-804A-5
14	128	97.0	2164	6	US-10-412-804A-5
15	128	97.0	2164	6	US-10-412-804A-5
16	128	97.0	2164	6	US-10-412-804A-5
17	128	97.0	2164	6	US-10-412-804A-5
18	128	97.0	2164	6	US-10-412-804A-5
19	128	97.0	2164	6	US-10-412-804A-5
20	128	97.0	2164	6	US-10-412-804A-5
21	128	97.0	2164	6	US-10-412-804A-5
22	128	97.0	2164	6	US-10-412-804A-5
23	128	97.0	2164	6	US-10-412-804A-5
24	128	97.0	2164	6	US-10-412-804A-5
25	128	97.0	2164	6	US-10-412-804A-5
26	128	97.0	2164	6	US-10-412-804A-5
27	128	97.0	2164	6	US-10-412-804A-5
28	128	97.0	2164	6	US-10-412-804A-5
29	128	97.0	2164	6	US-10-412-804A-5
30	128	97.0	2164	6	US-10-412-804A-5
31	128	97.0	2164	6	US-10-412-804A-5
32	128	97.0	2164	6	US-10-412-804A-5
33	128	97.0	2164	6	US-10-412-804A-5
34	128	97.0	2164	6	US-10-412-804A-5
35	128	97.0	2164	6	US-10-412-804A-5
36	128	97.0	2164	6	US-10-412-804A-5
37	128	97.0	2164	6	US-10-412-804A-5
38	128	97.0	2164	6	US-10-412-804A-5
39	128	97.0	2164	6	US-10-412-804A-5
40	128	97.0	2164	6	US-10-412-804A-5
41	128	97.0	2164	6	US-10-412-804A-5
42	128	97.0	2164	6	US-10-412-804A-5
43	128	97.0	2164	6	US-10-412-804A-5
44	128	97.0	2164	6	US-10-412-804A-5
45	128	97.0	2164	6	US-10-412-804A-5
46	128	97.0	2164	6	US-10-412-804A-5
47	128	97.0	2164	6	US-10-412-804A-5
48	128	97.0	2164	6	US-10-412-804A-5
49	128	97.0	2164	6	US-10-412-804A-5
50	128	97.0	2164	6	US-10-412-804A-5
51	128	97.0	2164	6	US-10-412-804A-5
52	128	97.0	2164	6	US-10-412-804A-5
53	128	97.0	2164	6	US-10-412-804A-5
54	128	97.0	2164	6	US-10-412-804A-5
55	114	86.4	3771	3	US-09-854-356-11
56	114	86.4	3952	3	US-10-949-667-9
57	114	86.4	3953	3	US-09-870-759-117
58	114	86.4	3955	3	US-09-854-356-10
59	114	86.4	3955	3	US-09-751-708A-117
60	114	86.4	3955	8	US-10-428-817A-113
61	114	86.4	3955	9	US-10-937-758A-94
62	109.5	83.0	3564	9	US-10-794-514A-331
63	108.5	82.2	3504	9	US-10-794-514A-329
64	59	44.7	61791	7	US-10-322-281-645
65	58	43.9	26345	5	US-10-087-192-1705
66	55	41.7	308	8	US-10-425-115-60865
67	55	41.7	541	8	US-10-425-115-103319
68	55	41.7	717	5	US-10-027-632-20298
69	55	41.7	717	5	US-10-027-632-20299
70	55	41.7	717	6	US-10-027-632-20298
71	55	41.7	717	6	US-10-027-632-20299
72	55	41.7	784	8	US-10-425-115-103315
73	55	41.7	39405	5	US-10-087-192-1285
74	54	40.9	1245	7	US-10-437-963-70738
75	54	40.9	1263	7	US-10-424-599-33137
76	53.5	40.5	1257	10	US-11-097-143-29378
77	53.5	40.5	4722	10	US-11-097-143-29377
78	53.5	40.5	21729	10	US-11-097-143-6472
79	53.5	40.5	27423	10	US-11-097-143-3643
80	53	40.2	70	6	US-10-388-360-117
81	53	40.2	70	8	US-10-758-307-52

Sequence 53, Appl  
Sequence 76, Appl  
Sequence 124, Appl  
Sequence 23, Appl  
Sequence 26, Appl  
Sequence 394, Ap  
Sequence 119394,  
Sequence 22332, A  
Sequence 41562, A  
Sequence 639836,  
Sequence 639837,  
Sequence 1563, Ap  
Sequence 1562, Ap  
Sequence 1650, Ap  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 1, Appl  
Sequence 9641, Ap  
Sequence 10, Appl

82 53 40.2 70 8 US-10-714-195-53  
83 53 40.2 70 9 US-10-883-303-76  
84 53 40.2 70 9 US-10-852-797-324  
85 53 40.2 70 9 US-10-872-063-23  
86 53 40.2 70 9 US-10-857-715-26  
87 53 40.2 578 8 US-10-021-323-9394  
88 53 40.2 993 7 US-10-425-115-119394  
89 53 40.2 1430 9 US-10-450-763-22332  
90 53 40.2 2432 6 US-10-369-493-41562  
91 52.5 39.8 468 4 US-09-925-065A-639836  
92 52.5 39.8 468 4 US-09-925-065A-639837  
93 52.5 39.8 1833 6 US-10-238-075-1563  
94 52.5 39.8 2748 6 US-10-238-075-1562  
95 52 39.4 456 3 US-09-864-761-1650  
96 52 39.4 471 3 US-09-978-303-10  
97 52 39.4 471 9 US-10-915-017-10  
98 52 39.4 639 8 US-10-765-466-1  
99 736 5 US-10-198-846-9641  
100 1318 9 US-10-691-532-10

ALIGNMENTS

RESULT 1

US-09-878-134-68  
; Sequence 68, Application US/09878134  
; Publication No. US20020086303A1  
; GENERAL INFORMATION:

; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: King, Gordon E.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; THERAPY AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.532  
; CURRENT APPLICATION NUMBER: US/09/878,134  
; CURRENT FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 377  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 68  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-878-134-68

Alignment Scores:  
Pred. No.: 1.59e-12 Length: 471  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-878-134-68 (1-471)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 65 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGACACAGCGGTGTGAGAAAGTGCAGC 124  
Qy 21 LysProCysAla 24  
Db 125 AAGCCCTGTGCC 136

RESULT 2

US-10-775-204-1475  
; Sequence 1475, Application US/10775204  
; Publication No. US20050186664A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Balance, David J.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins

; FILE REFERENCE: PF564  
; CURRENT APPLICATION NUMBER: US/10/775,204  
; CURRENT FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/341,811  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/360,000  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: 60/378,950  
; PRIOR FILING DATE: 2002-05-10  
; PRIOR APPLICATION NUMBER: 60/398,008  
; PRIOR FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: 60/411,355  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/414,984  
; PRIOR FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: 60/417,611  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/420,246  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: 60/423,623  
; PRIOR FILING DATE: 2002-11-05  
; PRIOR APPLICATION NUMBER: 60/351,360  
; PRIOR FILING DATE: 2002-01-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1475  
; LENGTH: 1260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-775-204-1475

Alignment Scores:  
Pred. No.: 4.88e-12 Length: 1260  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 9 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-775-204-1475 (1-1260)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 946 CCCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGACACAGCGGTGTGAGAAAGTGCAGC 1005  
Qy 21 LysProCysAla 24  
Db 1006 AAGCCCTGTGCC 1017

RESULT 3

US-10-956-373-25  
; Sequence 25, Application US/10956373  
; Publication No. US20050123538A1  
; GENERAL INFORMATION:  
; APPLICANT: Shemesh, Ronen  
; APPLICANT: Oren, Anat  
; APPLICANT: Rotman, Galit  
; APPLICANT: Sela-Tavor, Osnat  
; APPLICANT: Walach, Shira  
; APPLICANT: Sameah-Greenwald, Shirley  
; APPLICANT: Beiman, Merav  
; APPLICANT: Eshel, Dani  
; APPLICANT: Savitsky, Kinneret  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ErbB-2 POLYPEPTIDES AND KITS AND  
; METHODS USING SAME  
; FILE REFERENCE: 28399  
; CURRENT APPLICATION NUMBER: US/10/956,373  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 25  
; LENGTH: 1320



; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-956-373-25

Alignment Scores:  
Pred. No.: 5,14e-12 Length: 1320  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 9 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-956-373-25 (1-1320)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCTGACAAACCAAGAGGTGACAGCAGAGAGGTGAACACACGGTGTGAGAGTGCAGC 1155  
Qy 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167

## RESULT 4

US-10-956-373-3

; Sequence 3, Application US/10956373

; Publication No. US20050123538A1

; GENERAL INFORMATION:

; APPLICANT: Shemesh, Ronen

; APPLICANT: Oren, Anat

; APPLICANT: Rotman, Galit

; APPLICANT: Sela-Tavor, Osnat

; APPLICANT: Walach, Shira

; APPLICANT: Sameah-Greenwald, Shirley

; APPLICANT: Beiman, Merav

; APPLICANT: Eshel, Dani

; APPLICANT: Savitsky, Kinneret

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND

; FILE REFERENCE: METHODS USING SAME

; CURRENT APPLICATION NUMBER: US/10/956,373

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 3

; LENGTH: 1944

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-956-373-3

## Alignment Scores:

Pred. No.: 8e-12 Length: 1944  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 9 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-956-373-3 (1-1944)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCTGACAAACCAAGAGGTGACAGCAGAGAGGTGAACACACGGTGTGAGAGTGCAGC 1155  
Qy 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167

## RESULT 5

US-10-956-373-15

; Sequence 15, Application US/10956373

; Publication No. US20050123538A1

; GENERAL INFORMATION:

; APPLICANT: Shemesh, Ronen

; APPLICANT: Oren, Anat  
; APPLICANT: Rotman, Galit  
; APPLICANT: Sela-Tavor, Osnat  
; APPLICANT: Walach, Shira  
; APPLICANT: Sameah-Greenwald, Shirley  
; APPLICANT: Beiman, Merav  
; APPLICANT: Eshel, Dani  
; APPLICANT: Savitsky, Kinneret  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND

; FILE REFERENCE: METHODS USING SAME

; CURRENT APPLICATION NUMBER: US/10/956,373

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 15

; LENGTH: 2061

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-956-373-15

## Alignment Scores:

Pred. No.: 8.55e-12 Length: 2061  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 9 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-956-373-15 (1-2061)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCTGACAAACCAAGAGGTGACAGCAGAGAGGTGAACACACGGTGTGAGAGTGCAGC 1155  
Qy 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167

## RESULT 6

US-10-412-804A-3

; Sequence 3, Application US/10412804A

; Publication No. US20030228606A1

; GENERAL INFORMATION:

; APPLICANT: Jing, Shuqian

; APPLICANT: Tatarewicz, Suzanna

; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses

; FILE REFERENCE: Thereof

; FILE REFERENCE: 01-1624-A

; CURRENT APPLICATION NUMBER: US/10/412,804A

; PRIOR FILING DATE: 2003-04-11

; PRIOR FILING DATE: 2002-04-11

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2132

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (78)..(2132)

US-10-412-804A-3

## Alignment Scores:

Pred. No.: 8.89e-12 Length: 2132  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-412-804A-3 (1-2132)

```
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1023 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1082

QY 21 LysProCysAla 24
Db 1083 AAGCCCTGTGCC 1094

RESULT 7
US-10-412-804A-9
; Sequence 9, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Tatatewicz, Suzanna
; APPLICANT: Jing, Shuguan
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2145)
US-10-412-804A-9

Alignment Scores:
Pred. No.: 8,97e-12 Length: 2149
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
Gaps: 0
DB: 6

US-09-632-036F-6 (1-24) x US-10-412-804A-9 (1-2149)
; Sequence 5, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Tatatewicz, Suzanna
; APPLICANT: Jing, Shuguan
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2145)
US-10-412-804A-9

Alignment Scores:
Pred. No.: 8,97e-12 Length: 2149
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
Gaps: 0
DB: 6

US-09-632-036F-6 (1-24) x US-10-412-804A-9 (1-2149)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 8
US-10-412-804A-5
; Sequence 5, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Tatatewicz, Suzanna
; APPLICANT: Jing, Shuguan
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
US-10-412-804A-5

Alignment Scores:
Pred. No.: 9,79e-12 Length: 2320
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
Gaps: 0
DB: 9

US-09-632-036F-6 (1-24) x US-10-956-373-1 (1-2320)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 10
US-10-956-373-13
; Sequence 13, Application US/10956373
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
```

```
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
US-10-412-804A-5

Alignment Scores:
Pred. No.: 9,04e-12 Length: 2164
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
Gaps: 0
DB: 6

US-09-632-036F-6 (1-24) x US-10-412-804A-5 (1-2164)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 9
US-10-956-373-1
; Sequence 1, Application US/10956373
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
; APPLICANT: Oren, Anat
; APPLICANT: Rotman, Galit
; APPLICANT: Sela-Tavor, Osnat
; APPLICANT: Walach, Shira
; APPLICANT: Sameah-Greenwald, Shirley
; APPLICANT: Beiman, Merav
; APPLICANT: Eshel, Dani
; APPLICANT: Savitsky, Kinneret
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ErbB-2 POLYPEPTIDES AND KITS AND
; FILE REFERENCE: METHODS USING SAME
; FILE REFERENCE: 28399
; CURRENT APPLICATION NUMBER: US/10/956,373
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-373-1

Alignment Scores:
Pred. No.: 9,79e-12 Length: 2320
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
Gaps: 0
DB: 9

US-09-632-036F-6 (1-24) x US-10-956-373-1 (1-2320)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 10
US-10-956-373-13
; Sequence 13, Application US/10956373
; Publication No. US20050123538A1
; GENERAL INFORMATION:
; APPLICANT: Shemesh, Ronen
```

```
; APPLICANT: Oren, Anat
; APPLICANT: Rotman, Galit
; APPLICANT: Sela-Tavor, Osnat
; APPLICANT: Walach, Shira
; APPLICANT: Sameah-Greenwald, Shirley
; APPLICANT: Belman, Merav
; APPLICANT: Eshel, Dani
; APPLICANT: Savitsky, Kinneret
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL ERBB-2 POLYPEPTIDES AND KITS AND
; FILE REFERENCE: 28399
; CURRENT APPLICATION NUMBER: US/10/956,373
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2437
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-373-13
Alignment Scores:
Pred. No.: 1,04e-11 Length: 2437
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 9 Gaps: 0
US-09-632-036F-6 (1-24) x US-10-956-373-13 (1-2437)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACACCAAGAGGTGACAGCAGGATGGAAACACACACGCGTGTGAGAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167
RESULT 11
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5
Alignment Scores:
Pred. No.: 1,7e-11 Length: 3765
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0
```

```
US-09-632-036F-6 (1-24) x US-10-207-498-5 (1-3765)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACACCAAGAGGTGACAGCAGGATGGAAACACACACGCGTGTGAGAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017
RESULT 12
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwkowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8
Alignment Scores:
Pred. No.: 1,7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0
US-09-632-036F-6 (1-24) x US-09-811-123-8 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACACCAAGAGGTGACAGCAGGATGGAAACACACACGCGTGTGAGAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017
RESULT 13
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 1.7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-811-115-2 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGTCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 14
US-09-765-973-1
; Sequence 1, Application US/09765973
; Publication No. US2002039573A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND
; FILE OF INVENTION: TREATMENT OF HER-2/neu ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.496
; CURRENT FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-09-765-973-1

Alignment Scores:
Pred. No.: 1.7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-765-973-1 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGTCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 15
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
```

```
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 1.7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-854-356-9 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCTGTCACAAACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 16
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 1 7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-930-125-1 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACCAACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 17
US-09-984-092-3
; Sequence 3, Application US/09984092
; Publication No. US20040037840A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL THERAPEUTIC VACCINE FORMULATIONS
; FILE REFERENCE: P1011PC00
; CURRENT APPLICATION NUMBER: US/09/984,092
; CURRENT FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3768)
US-09-984-092-3

Alignment Scores:
Pred. No.: 1 7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x US-09-984-092-3 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACCAACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 18
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
```

```
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

Alignment Scores:
Pred. No.: 1 7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-313-644-1 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACCAACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 19
US-10-280-576-3
; Sequence 3, Application US/10280576
; Publication No. US20040044405A1
; GENERAL INFORMATION:
; APPLICANT: Wolff, Matthew R.
; TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
; FILE REFERENCE: 09820.189
; CURRENT APPLICATION NUMBER: US/10/280,576
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/343,732
; PRIOR FILING DATE: 2001-10-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-280-576-3

Alignment Scores:
Pred. No.: 1 7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-280-576-3 (1-3768)
Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACCAACCAAGAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1005
Qy 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 20
US-10-441-779C-3
; Sequence 3, Application US/10441779C
; Publication No. US20040141958A1
; GENERAL INFORMATION:
```

```
/ APPLICANT: Steinaa, Lucilla
/ APPLICANT: Mouritsen, Soren
/ APPLICANT: Gautam, Anand
/ APPLICANT: Haaning, Jesper
/ APPLICANT: Dalum, Iben
/ APPLICANT: Birk, Peter
/ APPLICANT: Leach, Dana
/ APPLICANT: Karlsson, Klaus
/ TITLE OF INVENTION: NOVEL METHODS FOR THERAPEUTIC VACCINATION
/ FILE REFERENCE: 4614-0116P
/ CURRENT APPLICATION NUMBER: US/10/441,779C
/ CURRENT FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: 09/413,186
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/105,011
/ PRIOR FILING DATE: 1998-10-20
/ PRIOR APPLICATION NUMBER: PA 1998 01361
/ PRIOR FILING DATE: 1998-10-05
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 3768
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(3768)
/ OTHER INFORMATION:
US-10-441-779C-3

Alignment Scores:
Pred. No.: 1-7e-11 Length: 3768
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: Gaps: 0

US-09-632-036F-6 (1-24) x US-10-441-779C-3 (1-3768)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 21
US-10-384-339C-52
/ Sequence 52, Application US/10384339C
/ Publication No. US20040175703A1
/ GENERAL INFORMATION:
/ APPLICANT: Kreutzer, Roland
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
/ FILE REFERENCE: 20200/2002
/ CURRENT APPLICATION NUMBER: US/10/384,339C
/ CURRENT FILING DATE: 2003-03-07
/ PRIOR APPLICATION NUMBER: PCT/EP02/00152
/ PRIOR FILING DATE: 2002-01-09
/ PRIOR APPLICATION NUMBER: DE 10100586.5
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: DE 10155280.7
/ PRIOR FILING DATE: 2001-10-26
/ PRIOR APPLICATION NUMBER: DE 10158411.3
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: DE 10160151.4
/ PRIOR FILING DATE: 2001-12-07
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 52
/ LENGTH: 3768

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 946 CCCCTGCACAAACAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1005

QY 21 LysProCysAla 24
Db 1006 AAGCCCTGTGCC 1017

RESULT 21
US-10-384-339C-52
/ Sequence 52, Application US/10384339C
/ Publication No. US20040175703A1
/ GENERAL INFORMATION:
/ APPLICANT: Kreutzer, Roland
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
/ FILE REFERENCE: 20200/2002
/ CURRENT APPLICATION NUMBER: US/10/384,339C
/ CURRENT FILING DATE: 2003-03-07
/ PRIOR APPLICATION NUMBER: PCT/EP02/00152
/ PRIOR FILING DATE: 2002-01-09
/ PRIOR APPLICATION NUMBER: DE 10100586.5
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: DE 10155280.7
/ PRIOR FILING DATE: 2001-10-26
/ PRIOR APPLICATION NUMBER: DE 10158411.3
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: DE 10160151.4
/ PRIOR FILING DATE: 2001-12-07
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 52
/ LENGTH: 3768

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1120 CCCCTGCACAAACAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAAGTGCAGC 1179

QY 21 LysProCysAla 24
Db 1180 AAGCCCTGTGCC 1191

RESULT 23
US-10-146-473-32
/ Sequence 32, Application US/10146473
/ Publication No. US2003010888A1
/ GENERAL INFORMATION:
/ APPLICANT: Scanlan, Matthew
/ APPLICANT: Gout, Ivan
/ APPLICANT: Stockert, Elisabeth
/ APPLICANT: Gure, Ali
```

```
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-146-473-32

Alignment Scores:
Pred. No.: 2,07e-11 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 5 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-146-473-32 (1-4473)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||
Db 1120 CCCTGCAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1179

Qy 21 LysProCysAla 24
|||
Db 1180 AAGCCCTGTGCC 1191

RESULT 24
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-207-655-44

Alignment Scores:
Pred. No.: 2,07e-11 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 5 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-207-655-44 (1-4473)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||
Db 1120 CCCTGCAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1179

Qy 21 LysProCysAla 24
|||
Db 1180 AAGCCCTGTGCC 1191

RESULT 25
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-207-655-44

Alignment Scores:
Pred. No.: 2,07e-11 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 5 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-207-655-44 (1-4473)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||
Db 1120 CCCTGCAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1179

Qy 21 LysProCysAla 24
|||
Db 1180 AAGCCCTGTGCC 1191

RESULT 26
; Sequence 5, Application US/10762128
; Publication No. US20040219161A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409C1
; CURRENT APPLICATION NUMBER: US/10/762,128
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 09/441,411
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-762-128-5

Alignment Scores:
Pred. No.: 2,07e-11 Length: 4473
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 8 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-762-128-5 (1-4473)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
|||
Db 1120 CCCTGCAACCAAGAGGTGACAGCAGAGGATGGAAACACACAGCGGTGTGAGAGTGCAGC 1179
```

```
QY      21 LysProCysAla 24
Db      1180 AAGCCCTGTGCC 1191

RESULT 27
US-10-723-860-8
; Sequence 8, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-8

Alignment Scores:
Pred. No.:      2,07e-11      Length:      4473
Score:          128.00      Matches:      23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%      Mismatches: 1
Query Match:    96.97%      Indels:      0
DB:             8      Gaps:      0

US-09-632-036F-6 (1-24) x US-10-723-860-8 (1-4473)

QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      1120 CCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1179

QY      21 LysProCysAla 24
Db      1180 AAGCCCTGTGCC 1191

RESULT 28
US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.:      2.1e-11      Length:      4530
Score:          128.00      Matches:      23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%      Mismatches: 1
Query Match:    96.97%      Indels:      0
DB:             3      Gaps:      0

QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      1096 CCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155

QY      21 LysProCysAla 24
```

```
US-09-632-036F-6 (1-24) x US-09-877-177-11 (1-4530)

QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      1096 CCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155

QY      21 LysProCysAla 24
Db      1156 AAGCCCTGTGCC 1167

RESULT 29
US-10-177-293-125
; Sequence 125, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-125

Alignment Scores:
Pred. No.:      2.1e-11      Length:      4530
Score:          128.00      Matches:      23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%      Mismatches: 1
Query Match:    96.97%      Indels:      0
DB:             6      Gaps:      0

US-09-632-036F-6 (1-24) x US-10-177-293-125 (1-4530)

QY      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      1096 CCCTGCACCAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGC 1155

QY      21 LysProCysAla 24
```



```

; OTHER INFORMATION:
US-10-338-730-1
Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-338-730-1 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20
Db 1096 CCCTGCAACCAAGAGGTGACAGCAGAGATGGAACACACGCGGTGTGAGAGTGCAGC 1155

RESULT 30
US-10-007-926A-119
; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119
Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-007-926A-119 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20
Db 1096 CCCTGCAACCAAGAGGTGACAGCAGAGATGGAACACACGCGGTGTGAGAGTGCAGC 1155

Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 31
US-10-338-730-1
; Sequence 1, Application US/10338730
; Publication No. US20030147905A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
; FILE REFERENCE: 5017C
; CURRENT APPLICATION NUMBER: US/10/338,730
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: US 09/527,487
; PRIOR FILING DATE: 2002-03-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)

```

```

; OTHER INFORMATION:
US-10-338-730-1
Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-338-730-1 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20
Db 1096 CCCTGCAACCAAGAGGTGACAGCAGAGATGGAACACACGCGGTGTGAGAGTGCAGC 1155

Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 32
US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124
Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-101-510-124 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCySer 20
Db 1096 CCCTGCAACCAAGAGGTGACAGCAGAGATGGAACACACGCGGTGTGAGAGTGCAGC 1155

Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 33
US-10-116-275-131
; Sequence 131, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors

```

```
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-131

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-116-275-131 (1-4530)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCTGCACAAACCAAGAGGTGACAGCAGAGGTGAACACACGCGGTGTGAGAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 34
US-10-426-836-11
; Sequence 11, Application US/10426836
; Publication No. US20030211530A1
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; FILE REFERENCE: 11220/169
; CURRENT APPLICATION NUMBER: US/10/426,836
; CURRENT FILING DATE: 2003-05-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-426-836-11

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-426-836-11 (1-4530)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCTGCACAAACCAAGAGGTGACAGCAGAGGTGAACACACGCGGTGTGAGAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 35
US-10-272-437A-27
; Sequence 27, Application US/10272437A
; Publication No. US20030216309A1
; GENERAL INFORMATION:
; APPLICANT: Krag, David N.
```

```
; APPLICANT: Pero, Stephanie C.
; APPLICANT: Oligino, Lyn
; TITLE OF INVENTION: BINDING PEPTIDES SPECIFIC FOR THE EXTRACELLULAR DOMAIN OF ERBB2
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: V00139.70056.US
; CURRENT APPLICATION NUMBER: US/10/272,437A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/329,183
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-272-437A-27

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-272-437A-27 (1-4530)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCTGCACAAACCAAGAGGTGACAGCAGAGGTGAACACACGCGGTGTGAGAAGTGCAGC 1155

QY 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 36
US-10-117-937-595
; Sequence 595, Application US/10117937
; Publication No. US2003022039A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPIOTOPE SEQUENCES
; FILE REFERENCE: CTLIMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 595
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-937-595

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-117-937-595 (1-4530)
```

```
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 37
US-10-392-113-45
; Sequence 45, Application US/10392113
; Publication No. US20030224993A1
; GENERAL INFORMATION:
; APPLICANT: Land, Hartmut
; APPLICANT: Deleu, Laurent
; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION
; TITLE OF INVENTION: OF CANCER CELLS
; FILE REFERENCE: 21108.0005U3
; CURRENT APPLICATION NUMBER: US/10/392,113
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/365,078
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: PCT/US01/32127
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/239,705
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: Synthetic Construct
US-10-392-113-45

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-392-113-45 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 38
US-10-159-563-208
; Sequence 208, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 208
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-208

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-159-563-208 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 39
US-10-435-696-10
; Sequence 10, Application US/10435696
; Publication No. US20040018525A1
; GENERAL INFORMATION:
; APPLICANT: Wirtz, Ralph
; APPLICANT: Munnes, Marc
; APPLICANT: Kallabis, Harald
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA
; FILE REFERENCE: Lea 36 108
; CURRENT APPLICATION NUMBER: US/10/435,696
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: EP03003112.4
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: EP02010291.9
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-435-696-10

Alignment Scores:
Pred. No.: 2,1e-11 Length: 4530
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-435-696-10 (1-4530)
Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 1096 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAAGTGCAGC 1155
Qy 21 LysProCysAla 24
Db 1156 AAGCCCTGTGCC 1167

RESULT 40
US-10-734-564-59
; Sequence 59, Application US/10734564
; Publication No. US20040157278A1
; GENERAL INFORMATION:
; APPLICANT: Christopher C Burgess et al
```

; TITLE OF INVENTION: Detection Methods Using TIMP1

; FILE REFERENCE: 1657/2012

; CURRENT APPLICATION NUMBER: US/10/734,564

; CURRENT FILING DATE: 2003-12-12

; NUMBER OF SEQ ID NOS: 138

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 59

; LENGTH: 4530

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-734-564-59

Alignment Scores:

Pred. No.:	2,1e-11	Length:	4530
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	7	Gaps:	0

US-09-632-036F-6 (1-24) x US-10-734-564-59 (1-4530)

QY	1	ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer	20
DB	1096	CCCCCTGCACCAACCAAGAGGTGACAGAGGTGAACACAGCGGTGTGAGAAAGTGCAGC	1155
QY	21	LysProCysAla	24
DB	1156	AAGCCCTGTGCC	1167

Search completed: December 11, 2005, 03:11:47

491 182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 11, 2005, 01:00:20 ; Search time 143.455 Seconds  
(without alignment)  
62.546 Million cell updates/sec

Title: US-09-632-036F-6

Perfect score: 132

Sequence: 1 PLNNQVTAEDGTQRAEKSCKPCA 24

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09632036/runat\_02122005\_104143\_25959/app\_query.fasta\_1.398  
-DB=Published Applications\_NA\_New -QFMT=fastcap -SUFFIX=rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62  
-TRANS=human40.cdi -LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -THRS=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09632036@cgn.1.1.227@runat\_02122005\_104143\_25959  
-NCPU=6 -ICPU=1 -NO MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3:  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3:  
10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	4530	6	US-10-770-726-18
2	128	97.0	149419	7	US-11-112-908-49
3	128	97.0	157224	7	US-11-112-908-51
4	128	97.0	161726	7	US-11-112-908-48
5	128	97.0	161726	7	US-11-112-908-52
C 6	57	43.2	6590	6	US-10-467-657-3887
7	54	40.9	184868	7	US-11-121-086-88
C 8	51.5	39.0	1302	6	US-10-467-657-345
					Sequence 18, Appl
					Sequence 49, Appl
					Sequence 51, Appl
					Sequence 48, Appl
					Sequence 52, Appl
					Sequence 3887, Ap
					Sequence 88, Appl
					Sequence 345, App

C 9	51	38.6	1763	6	US-10-632-150-13	Sequence 13, Appl
C 10	51	38.6	1763	7	US-11-073-457-13	Sequence 13, Appl
C 11	51	38.6	1794	7	US-11-074-176-257	Sequence 257, App
C 12	50.5	38.3	780	9	US-11-082-389-249	Sequence 249, App
C 13	49.5	37.5	2766	6	US-10-750-185-55939	Sequence 55939, A
C 14	49.5	37.5	178877	7	US-11-121-086-17	Sequence 17, Appl
C 15	49	37.1	600	6	US-10-750-185-1366	Sequence 1366, Ap
C 16	49	37.1	168516	7	US-11-121-086-3	Sequence 3, Appli
C 17	49	37.1	190882	7	US-11-121-086-69	Sequence 69, Appl
C 18	48	36.4	1430	6	US-10-750-185-61160	Sequence 61160, A
C 19	48	36.4	1795	6	US-10-750-185-41622	Sequence 41622, A
C 20	48	36.4	124972	7	US-11-121-086-100	Sequence 100, App
C 21	48	36.4	134174	7	US-11-121-086-99	Sequence 99, Appl
C 22	47	35.6	1634	6	US-10-750-185-40234	Sequence 40234, A
C 23	47	35.6	1082144	7	US-11-117-187-211	Sequence 211, App
C 24	46.5	35.2	1219	6	US-10-750-185-54567	Sequence 54567, A
C 25	46	34.8	1490	6	US-10-750-185-47299	Sequence 47299, A
C 26	46	34.8	1670	6	US-10-614-599-2	Sequence 2, Appli
C 27	46	34.8	2151	7	US-11-147-047-26	Sequence 26, Appl
C 28	46	34.8	2254	6	US-10-750-185-56790	Sequence 56790, A
C 29	46	34.8	2377	6	US-10-750-185-62314	Sequence 62314, A
C 30	46	34.8	2658	6	US-10-131-826A-511	Sequence 511, App
C 31	46	34.8	64415	7	US-11-117-187-185	Sequence 185, App
C 32	46	34.8	150314	7	US-11-112-908-24	Sequence 24, Appl
C 33	46	34.8	163317	7	US-11-117-187-212	Sequence 212, App
C 34	46	34.8	189539	7	US-11-121-086-16	Sequence 16, Appl
C 35	46	34.8	218821	7	US-11-121-086-31	Sequence 31, Appl
C 36	45.5	34.5	153142	7	US-11-121-086-27	Sequence 27, Appl
C 37	45	34.1	1649	6	US-10-750-185-24998	Sequence 24998, A
C 38	45	34.1	1713	6	US-10-750-185-56624	Sequence 56624, A
C 39	45	34.1	2838	6	US-10-750-185-46440	Sequence 46440, A
C 40	45	34.1	92584	7	US-11-117-187-194	Sequence 194, App
C 41	45	34.1	150314	7	US-11-112-908-24	Sequence 24, Appl
C 42	45	34.1	180654	7	US-11-121-086-58	Sequence 58, Appl
C 43	44.5	33.7	816	6	US-10-750-185-60284	Sequence 60284, A
C 44	44.5	33.7	1609	6	US-10-750-185-55040	Sequence 55040, A
C 45	44.5	33.7	3334	6	US-10-750-185-34055	Sequence 34055, A
C 46	44.5	33.7	167891	7	US-11-121-086-14	Sequence 14, Appl
C 47	44	33.3	1056	6	US-10-467-657-2437	Sequence 2437, Ap
C 48	44	33.3	1286	6	US-10-750-185-25304	Sequence 25304, A
C 49	44	33.3	1321	6	US-10-750-185-49780	Sequence 49780, A
C 50	44	33.3	1356	6	US-10-750-185-53605	Sequence 53605, A
C 51	44	33.3	2819	8	US-11-112-944-13	Sequence 13, Appl
C 52	44	33.3	3432	6	US-10-467-657-1819	Sequence 1819, Ap
C 53	44	33.3	3602	6	US-10-750-185-28825	Sequence 28825, A
C 54	44	33.3	79122	7	US-11-117-187-200	Sequence 200, App
C 55	44	33.3	148220	7	US-11-121-086-90	Sequence 90, Appl
C 56	44	33.3	150481	7	US-11-112-908-37	Sequence 37, Appl
C 57	44	33.3	179777	7	US-11-121-086-106	Sequence 106, App
C 58	44	33.3	179892	7	US-11-112-908-39	Sequence 39, Appl
C 59	44	33.3	197096	7	US-11-121-086-107	Sequence 107, App
C 60	43.5	33.0	600	6	US-10-750-185-21076	Sequence 21076, A
C 61	43.5	33.0	1124	6	US-10-750-185-41659	Sequence 41659, A
C 62	43.5	33.0	1295	6	US-10-131-826A-335	Sequence 335, App
C 63	43.5	33.0	185393	7	US-11-121-086-101	Sequence 101, App
C 64	43	32.6	600	6	US-10-750-185-21492	Sequence 21492, A
C 65	43	32.6	600	6	US-10-750-185-41507	Sequence 41507, A
C 66	43	32.6	670	6	US-10-750-185-61723	Sequence 61723, A
C 67	43	32.6	1340	6	US-10-750-185-57537	Sequence 57537, A
C 68	43	32.6	1377	6	US-10-742-634-6	Sequence 6, Appli
C 69	43	32.6	1561	6	US-10-750-185-48239	Sequence 48239, A
C 70	43	32.6	1580	6	US-10-750-185-32268	Sequence 32268, A
C 71	43	32.6	1612	6	US-10-750-185-31631	Sequence 31631, A
C 72	43	32.6	1803	6	US-10-750-185-40283	Sequence 40283, A
C 73	43	32.6	1869	6	US-10-750-185-48903	Sequence 48903, A
C 74	43	32.6	2326	6	US-10-750-185-38395	Sequence 38395, A
C 75	43	32.6	2721	6	US-10-750-185-52450	Sequence 52450, A
C 76	43	32.6	3591	6	US-10-667-295-99	Sequence 99, Appl
C 77	43	32.6	5260	6	US-10-750-185-25097	Sequence 25097, A
C 78	43	32.6	5796	6	US-10-821-234-62	Sequence 62, Appl
C 79	43	32.6	6087	7	US-11-000-463-485	Sequence 485, App
C 80	43	32.6	6116	7	US-11-000-463-13	Sequence 13, Appl
C 81	43	32.6	154548	7	US-11-121-086-33	Sequence 33, Appl

Sequence 43, Appl  
Sequence 20, Appl  
Sequence 50, Appl  
Sequence 24, Appl  
Sequence 36, Appl  
Sequence 50, Appl  
Sequence 11, Appl  
Sequence 29939, A  
Sequence 32500, A  
Sequence 7, Appl  
Sequence 96, Appl  
Sequence 51, Appl  
Sequence 77, Appl  
Sequence 401, Appl  
Sequence 923, Appl  
Sequence 230, Appl  
Sequence 28491, A  
Sequence 31610, A  
Sequence 37352, A

ALIGNMENTS

US-10-770-726-18  
; Sequence 18, Application US/10770726  
; Publication No. US20050266409A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Brown, Eugene  
; APPLICANT: Liu, Wei  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; FILE REFERENCE: AM101079 (031896-010000)  
; CURRENT FILING DATE: 2004-02-04  
; NUMBER OF SEQ ID NOS: 48640  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-770-726-18

Alignment Scores:  
Pred. No.: 1.68e-11 Length: 4530  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservatives: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-770-726-18 (1-4530)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1096 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGC 1155  
QY 21 LysProCysAla 24  
Db 1156 AAGCCCTGTGCC 1167

; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 49  
; LENGTH: 149419  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-49

Alignment Scores:  
Pred. No.: 1.26e-09 Length: 149419  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservatives: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-112-908-49 (1-149419)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 6108 CCCCTGCACAAACCAAGAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGC 6167  
QY 21 LysProCysAla 24  
Db 6168 AAGCCCTGTGCC 6179

Alignment Scores:  
Pred. No.: 1.34e-09 Length: 157224  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservatives: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-112-908-51 (1-157224)  
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20

Db 142585 CCCTGCAACCAAGAGGTGCACAGAGGATGGAAACACACCGGTGTGTGAGAAGTGCAGC 142644

Qy 21 LysProCysAla 24  
|||||

Db 142645 AAGCCCTGTGCC 142656

RESULT 4

US-11-112-908-48

; Sequence 48, Application US/11112908

; Publication No. US20050260659A1

; GENERAL INFORMATION:

; APPLICANT: Harris, Cole

; APPLICANT: Davis, Lisa M.

; TITLE OF INVENTION: Breast Cancer Biomarkers

; FILE REFERENCE: 04-164-US

; CURRENT APPLICATION NUMBER: US/11/112,908

; CURRENT FILING DATE: 2005-04-22

; PRIOR APPLICATION NUMBER: US 60/564,758

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/575,978

; PRIOR FILING DATE: 2004-06-01

; PRIOR APPLICATION NUMBER: US 60/631,702

; PRIOR FILING DATE: 2004-11-30

; PRIOR APPLICATION NUMBER: US 60/633,826

; PRIOR FILING DATE: 2004-12-07

; NUMBER OF SEQ ID NOS: 511

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 48

; LENGTH: 161726

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-112-908-48

Alignment Scores:

Pred. No.:	1.39e-09	Length:	161726
Score:	128.00	Matches:	23
Percent Similarity:	95.83%	Conservative:	0
Best Local Similarity:	95.83%	Mismatches:	1
Query Match:	96.97%	Indels:	0
DB:	7	Gaps:	0

US-09-632-036F-6 (1-24) x US-11-112-908-48 (1-161726)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAaspGlyThrGlnArgAlaGluLysCysSer 20  
|||||

Db 56390 CCCTGCAACCAAGAGGTGCACAGAGGATGGAAACACACCGGTGTGTGAGAAGTGCAGC 56449

Qy 21 LysProCysAla 24  
|||||

Db 56450 AAGCCCTGTGCC 56461

RESULT 5

US-11-112-908-52

; Sequence 52, Application US/11112908

; Publication No. US20050260659A1

; GENERAL INFORMATION:

; APPLICANT: Harris, Cole

; APPLICANT: Davis, Lisa M.

; TITLE OF INVENTION: Breast Cancer Biomarkers

; FILE REFERENCE: 04-164-US

; CURRENT APPLICATION NUMBER: US/11/112,908

; CURRENT FILING DATE: 2005-04-22

; PRIOR APPLICATION NUMBER: US 60/564,758

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/575,978

; PRIOR FILING DATE: 2004-06-01

; PRIOR APPLICATION NUMBER: US 60/631,702

; PRIOR FILING DATE: 2004-11-30

; PRIOR APPLICATION NUMBER: US 60/633,826

; PRIOR FILING DATE: 2004-12-07

; NUMBER OF SEQ ID NOS: 511

; SOFTWARE: PatentIn version 3.3

SEQ ID NO 52

```

; LENGTH: 161726
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-52

Alignment Scores:
Pred. No.: 1.39e-09 Length: 161726
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-112-908-52 (1-161726)

Qy 1 ProLeuHisAsnGlnCluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 56390 CCCTGCACAACCAAGAGGTGACAGCAGAGGATGGAAACACACACGCGGTGTGAGAAAGTGCAGC 56449

Qy 21 LysProCysAla 24
Db 56450 AAGCCCTGTGCC 56461

RESULT 6
US-10-467-657-3887/c
; Sequence 3887, Application US10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3887
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3887

Alignment Scores:
Pred. No.: 3.62 Length: 690
Score: 57.00 Matches: 11
Percent Similarity: 63.64% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 43.18% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-467-657-3887 (1-690)

Qy 3 HisAsnGlnCluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysPro 22
Db 600 AATAACCAACGACGCGCGATGCCGAACCAACGCGCGATGCGCGCGCCCAAGAGATAAAGCCAAA 541

Qy 23 CysAla 24
Db 540 TGTGCG 535

RESULT 7
US-11-121-086-88
; Sequence 88, Application US/11121086
; Publication No. US2005026459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

```

```

; LENGTH: 161726
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-52

Alignment Scores:
Pred. No.: 1.39e-09 Length: 161726
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-112-908-52 (1-161726)

Qy 1 ProteuHisenGlnGluValThrAlaGluApsGlyThrGlnArgAlaGluLysCysSer 20
Db 56390 CCCCTGTCACCAACCAAGAGGTGCACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGCAGC 564
Qy 21 LysProCysAla 24
Db 56450 AAGCCCTGTGCC 56461

RESULT 6
US-10-467-657-3887/c
; Sequence 3887, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNIGANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467.657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3887
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3887

Alignment Scores:
Pred. No.: 3.62 Length: 690
Score: 57.00 Matches: 11
Percent Similarity: 63.64% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 43.18% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-467-657-3887 (1-690)

Qy 3 HisAsnGlnGluValThrAlaGluApsGlyThrGlnArgAlaGluLysCysSerLysPro 22
Db 600 AATAACACGACACGCGATGTCGGAACCGGCAGTCAGCGCGCCCAAGATAAAGCAAAA 541
Qy 23 CysAla 24
Db 540 TGTGCG 535

RESULT 7
US-11-121-086-88
; Sequence 88, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

```

FILE REFERENCE: 09138.6000-00000  
CURRENT APPLICATION NUMBER: US/11/121,086  
CURRENT FILING DATE: 2005-05-04  
PRIOR APPLICATION NUMBER: 60/567,570  
PRIOR FILING DATE: 2004-05-04  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: Patent in version 3.3  
SEQ ID NO 88  
LENGTH: 184868  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-121-086-88

Alignment Scores:  
Pred. No.: 1,2e+04 Length: 184868  
Score: 54.00 Matches: 11  
Percent Similarity: 65.00% Conservative: 2  
Best Local Similarity: 55.00% Mismatches: 7  
Query Match: 40.91% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-121-086-88 (1-184868)

Qy 5 GlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24  
Db 1665 CAGGAGCTCACAGCTGAGATGGGGTTTCAGAAATGAAGTAAGTAAAGCCCGCAGTGTGCC 1724

## RESULT 8

US-10-467-657-345/c  
Sequence 345, Application US/10467657  
Publication No. US20050260581A1

## GENERAL INFORMATION:

APPLICANT: CHIRON SpA  
APPLICANT: FONTANA Maria Rita  
APPLICANT: PIZZA Mariagrazia  
APPLICANT: MASIGNANI Vega  
APPLICANT: MONACI Elisabetta  
TITLE OF INVENTION: GONGOCOCCAL PROTEINS AND NUCLEIC ACIDS  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/467,657  
CURRENT FILING DATE: 2003-08-11  
PRIOR APPLICATION NUMBER: GB-0103424.8  
PRIOR FILING DATE: 2001-02-12  
NUMBER OF SEQ ID NOS: 9218  
SOFTWARE: SeqWin99, version 1.04  
SEQ ID NO 345  
LENGTH: 1302  
TYPE: DNA  
ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-345

Alignment Scores:  
Pred. No.: 71.7 Length: 1302  
Score: 51.50 Matches: 10  
Percent Similarity: 57.69% Conservative: 5  
Best Local Similarity: 38.46% Mismatches: 8  
Query Match: 39.02% Indels: 3  
DB: 6 Gaps: 1

US-09-632-036F-6 (1-24) x US-10-467-657-345 (1-1302)

Qy 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArg-----AlaGlu 17  
Db 816 CCCATCCACATCAGCGTATCGGCCTGATTGACGGGAATCAGCGGAATCTCGGCCGGAAG 757

Qy 18 LysCysSerLysProCys 23  
Db 756 CGACCTCTCTCGACCGTGC 739

## RESULT 9

US-10-632-150-13/c  
Sequence 13, Application US/10632150  
Publication No. US20050251871A1

GENERAL INFORMATION:  
APPLICANT: Chiau, D.  
APPLICANT: Pagano, M.  
APPLICANT: Latres, E.  
TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS  
FILE REFERENCE: 5914-081  
CURRENT APPLICATION NUMBER: US/10/632,150  
CURRENT FILING DATE: 2003-07-30  
PRIOR APPLICATION NUMBER: US/09/385,219  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: 60/098,355  
PRIOR FILING DATE: 1998-08-28  
PRIOR APPLICATION NUMBER: 60/118,568  
PRIOR FILING DATE: 1999-02-03  
PRIOR APPLICATION NUMBER: 60/124,449  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 13  
LENGTH: 1763  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-632-150-13

Alignment Scores:  
Pred. No.: 127 Length: 1763  
Score: 51.00 Matches: 9  
Percent Similarity: 72.22% Conservative: 4  
Best Local Similarity: 50.00% Mismatches: 5  
Query Match: 38.64% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-632-150-13 (1-1763)

Qy 3 HisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1567 CACAATCAGAAATAACACTCGAGATCAGCACCACCAAGGAGTGACATCTGTAGT 1514

## RESULT 10

US-11-073-457-13/c  
Sequence 13, Application US/11073457  
Publication No. US20050260556A1

## GENERAL INFORMATION:

APPLICANT: Pagano, M.  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIF  
FILE REFERENCE: 5914-090-999  
CURRENT APPLICATION NUMBER: US/11/073,457  
CURRENT FILING DATE: 2005-03-04  
PRIOR APPLICATION NUMBER: 10/042,417  
PRIOR FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 13  
LENGTH: 1763  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-073-457-13

Alignment Scores:  
Pred. No.: 127 Length: 1763  
Score: 51.00 Matches: 9  
Percent Similarity: 72.22% Conservative: 4  
Best Local Similarity: 50.00% Mismatches: 5  
Query Match: 38.64% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-073-457-13 (1-1763)

Qy 3 HisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 1567 CACAATCAGAAATAACACTCGAGATCAGCACCACCAAGGAGTGACATCTGTAGT 1514



```
Db      1567 CACAAATCAGAAATAAACAACACTCGAGATCAGACACCCCAAGGAGTGACATCTGTAGT 1514
RESULT 11
US-11-074-176-257/c
; Sequence 257, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Aacarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257
; LENGTH: 1794
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1794)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 1763; oligopeptidase
US-11-074-176-257
Alignment Scores:
Pred. No.:      130      Length:      1794
Score:          51.00    Matches:      13
Percent Similarity: 68.18% Conservative: 2
Best Local Similarity: 59.09% Mismatches: 5
Query Match:     38.64% Indels:      2
DB:              7      Gaps:        2

US-09-632-036f-6 (1-24) x US-11-074-176-257 (1-1794)
Qy      2 LeuHieaenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGlu---LysCysSer 20
Db      859 CTACATAACCGTGAAGTAAT---CAAGATGGGAGTCAACTTTCGAATCAAAAGTGTGCT 803
Qy      21 LysPro 22
Db      802 AAACCT 797

RESULT 12
US-11-082-389-249/c
; Sequence 249, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habethauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262

US-09-632-036f-6 (1-24) x US-11-082-389-249 (1-780)
Qy      6 GluValThrAlaGluAspGlyThrGlnArgAlaGluLys-----CysSerLysPro 22
Db      585 GAGTCCACGGCGCGGAGGGGTTCTCGAGAAGCAATAAGGGCGCCTGTTGGCTAAACCT 526
Qy      23 CysAla 24
Db      525 TGGCT 520

RESULT 13
US-10-750-185-55939
; Sequence 55939, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 55939
; LENGTH: 2766
; TYPE: DNA
; ORGANISM: Bovine 1986688107767
US-10-750-185-55939
Alignment Scores:
Pred. No.:      405      Length:      2766
```

Score: 49.50 Matches: 13  
Percent Similarity: 69.57% Conservative: 3  
Best Local Similarity: 56.52% Mismatches: 5  
Query Match: 37.50% Indels: 2  
DB: 6 Gaps: 1

US-09-632-036F-6 (1-24) x US-10-750-185-55939 (1-2766)

QY 1 ProLeuHisAenGlnGluValThrAlaGlu-AspGlyThrGlnArgAlaGluLysCysSe 20  
Db 2067 CCACCTTCAGACAGTGAA---ACGCCAACCAAGGACGAGGAGGTAGCTGTGTTCC 2123

QY 20 rLysPro 22  
Db 2124 AAAGCCA 2130

RESULT 14  
US-11-121-086-17  
; Sequence 17, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121.086  
; CURRENT FILING DATE: 2005-05-04  
; PRIOR FILING DATE: 2004-05-04  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 17  
; LENGTH: 178877  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-17

Alignment Scores:  
Pred. No.: 6.93e+04 Length: 178877  
Score: 49.50 Matches: 9  
Percent Similarity: 71.43% Conservative: 6  
Best Local Similarity: 42.86% Mismatches: 5  
Query Match: 37.50% Indels: 1  
DB: 7 Gaps: 1

US-09-632-036F-6 (1-24) x US-11-121-086-17 (1-178877)

QY 4 AenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCys 23  
Db 96206 AACAGGCAACTCACT---CAAGAAGGAATACACAGGGCCAATAAATGTATGAAAGTTGC 96262

QY 24 Ala 24  
Db 96263 TCA 96265

RESULT 15  
US-10-750-185-1366/c  
; Sequence 1366, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750.185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31

Score: 49.50 Matches: 13  
Percent Similarity: 69.57% Conservative: 3  
Best Local Similarity: 56.52% Mismatches: 5  
Query Match: 37.50% Indels: 2  
DB: 6 Gaps: 1

US-09-632-036F-6 (1-24) x US-10-750-185-55939 (1-2766)

QY 1 ProLeuHisAenGlnGluValThrAlaGlu-AspGlyThrGlnArgAlaGluLysCysSe 20  
Db 2067 CCACCTTCAGACAGTGAA---ACGCCAACCAAGGACGAGGAGGTAGCTGTGTTCC 2123

QY 20 rLysPro 22  
Db 2124 AAAGCCA 2130

RESULT 14  
US-11-121-086-17  
; Sequence 17, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121.086  
; CURRENT FILING DATE: 2005-05-04  
; PRIOR FILING DATE: 2004-05-04  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 17  
; LENGTH: 178877  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-17

Alignment Scores:  
Pred. No.: 6.93e+04 Length: 178877  
Score: 49.50 Matches: 9  
Percent Similarity: 71.43% Conservative: 6  
Best Local Similarity: 42.86% Mismatches: 5  
Query Match: 37.50% Indels: 1  
DB: 7 Gaps: 1

US-09-632-036F-6 (1-24) x US-11-121-086-17 (1-178877)

QY 4 AenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCys 23  
Db 96206 AACAGGCAACTCACT---CAAGAAGGAATACACAGGGCCAATAAATGTATGAAAGTTGC 96262

QY 24 Ala 24  
Db 96263 TCA 96265

RESULT 15  
US-10-750-185-1366/c  
; Sequence 1366, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750.185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31

```
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69
; LENGTH: 190882
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-69

Alignment Scores:
Pred. No.:          9,16e+04      Length:      190882
Score:              49.00         Matches:      9
Percent Similarity: 60.87%        Conservative: 5
Best Local Similarity: 39.13%     Mismatches:  5
Query Match:        37.12%        Indels:      0
DB:                 7            Gaps:         0

US-09-632-036F-6 (1-24) x US-11-121-086-69 (1-190882)

Qy      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
      |||||
Db 189754 CCAGTGCACCTCCAGCGCTGGTGACAGAGGAGATTCTCTCAAAAAAAAAAGAGATGCGCTG 189695

Qy      21 LysProCys 23
      |||
Db 189694 TGTCTGTGC 189686

RESULT 18
US-10-750-185-61160
; Sequence 61160, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61160
; LENGTH: 1430
; TYPE: DNA
; ORGANISM: Bovine 19866880738483
US-10-750-185-61160

Alignment Scores:
Pred. No.:          327          Length:      1430
Score:              48.00         Matches:      8
Percent Similarity: 66.67%        Conservative:  4
Best Local Similarity: 44.44%     Mismatches:  6
Query Match:        36.36%        Indels:      0
DB:                 6            Gaps:         0

US-09-632-036F-6 (1-24) x US-10-750-185-61160 (1-1430)

Qy      7 ValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24
      |||||
Db 374 GTCTGTCGACGATGGGAGCTACAAAGCTGAAAGAGCTCCACAGCCCTGCTCC 427

RESULT 19
US-10-750-185-41622
; Sequence 41622, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41622
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Bovine 19866880973184
US-10-750-185-41622

Alignment Scores:
Pred. No.:          432          Length:      1795
Score:              48.00         Matches:      9
Percent Similarity: 60.00%        Conservative:  3
Best Local Similarity: 45.00%     Mismatches:  8
Query Match:        36.36%        Indels:      0
DB:                 6            Gaps:         0

US-09-632-036F-6 (1-24) x US-10-750-185-41622 (1-1795)

Qy      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
      |||||
Db 189 CCCCCCACCACCCAGCCACCCACCCAGGTCAATGGAGCTCTAGCCAGGAGAAATGTGCA 248

RESULT 20
US-11-121-086-100/c
; Sequence 100, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 100
; LENGTH: 124972
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-100

Alignment Scores:
Pred. No.:          8,11e+04      Length:      124972
Score:              48.00         Matches:      9
Percent Similarity: 52.17%        Conservative:  3
Best Local Similarity: 39.13%     Mismatches: 11
Query Match:        36.36%        Indels:      0
DB:                 7            Gaps:         0

US-09-632-036F-6 (1-24) x US-11-121-086-100 (1-124972)

Qy      1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
      |||||
Db 4340 CCAGTGCATGACCAAGAAATATATATGGCGCTGAGACAGCTGAGACAAGGTGAGGCATGCGCA 4281
```

QY 21 LysProCys 23  
|||  
Db 4280 GGCCAATGT 4272

RESULT 21  
US-11-121-086 99/c  
; Sequence 99, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; CURRENT FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 99  
; LENGTH: 134174  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-99

Alignment Scores:  
Pred. No.: 8.85e+04 Length: 134174  
Score: 48.00 Matches: 9  
Percent Similarity: 52.17% Conservative: 3  
Best Local Similarity: 39.13% Mismatches: 11  
Query Match: 36.36% Indels: 0  
Gaps: 0  
DB:

US-09-632-036F-6 (1-24) x US-11-121-086-99 (1-134174)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 109403 CCAGTGCATGACCAAGAATATATGGCTGAGACAGCTGAGACAAAGGGTGAGGCATGCCAA 109344

QY 21 LysProCys 23  
|||  
Db 109343 GGCCAATGT 109335

RESULT 22  
US-10-750-185-40234  
; Sequence 40234, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40234  
; LENGTH: 1634  
; TYPE: DNA  
; ORGANISM: Bovine 19866881082730  
US-10-750-185-40234

Alignment Scores:  
Pred. No.: 575 Length: 1634  
Score: 47.00 Matches: 10  
Percent Similarity: 68.75% Conservative: 1

Best Local Similarity: 62.50% Mismatches: 5  
Query Match: 35.61% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-750-185-40234 (1-1634)

QY 2 LeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGlu 17  
|||  
Db 541 CTGCACAGACGAGAAGTGAAGGCAGAGGATAACTGCTGAGGCTGAA 588

RESULT 23

US-11-117-187-211  
; Sequence 211, Application US/11117187  
; Publication No. US20050266560A1  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVER, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARCD:309US  
; CURRENT APPLICATION NUMBER: US/11/117,187  
; CURRENT FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: US/09/531,120  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 211  
; LENGTH: 1082144  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-11-117-187-211

Alignment Scores:  
Pred. No.: 1.54e+06 Length: 1082144  
Score: 47.00 Matches: 10  
Percent Similarity: 47.83% Conservative: 1  
Best Local Similarity: 43.48% Mismatches: 12  
Query Match: 35.61% Indels: 0  
Gaps: 0  
DB:

US-09-632-036F-6 (1-24) x US-11-117-187-211 (1-1082144)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||  
Db 860624 CCTTTTCAACGACACATCGACCGCATAGCTCGGAAGTCTGAGAAAGGAGAAATACAGT 860683

QY 21 LysProCys 23  
|||  
Db 860684 CAACCTTGT 860692

RESULT 24

US-10-750-185-54567  
; Sequence 54567, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54567  
; LENGTH: 1219





```
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 511
; LENGTH: 2668
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-511

Alignment Scores:
Pred. No.: 1.57e+03 Length: 2668
Score: 46.00 Matches: 9
Percent Similarity: 50.00% Conservative: 2
Best Local Similarity: 40.91% Mismatches: 11
Query Match: 34.85% Indels: 0
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x US-10-131-826A-511 (1-2668)

Qy 2 LeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLys 21
   :::::
Db 2067 GTACACGACGATGGACGAGGCTTAATGACGGCAACATGACCCCTTACTGCAGCAAG 2008
   :::::

Qy 22 ProCys 23
   :::::
Db 2007 GGCTGT 2002

RESULT 31
US-11-117-187-185
; Sequence 185, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 64415
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (9960)..(21146)
; OTHER INFORMATION: N = A, C G, or T/U
US-11-117-187-185

Alignment Scores:
Pred. No.: 7.97e+04 Length: 64415
Score: 46.00 Matches: 10
Percent Similarity: 47.83% Conservative: 1
Best Local Similarity: 43.48% Mismatches: 12
Query Match: 34.85% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-117-187-185 (1-64415)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||
Db 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||
```

```
Db 56937 CCTTTTCAACGACAAATCGACCGCATAGCTCGGAACTCAGAGAAAGGAAAAACACAGT 56996

Qy 21 LysProCys 23
   :::::
Db 56997 CAACCTTGT 57005

RESULT 32
US-11-112-908-24/c
; Sequence 24, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 150314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-24

Alignment Scores:
Pred. No.: 2.24e+05 Length: 150314
Score: 46.00 Matches: 9
Percent Similarity: 61.90% Conservative: 4
Best Local Similarity: 42.86% Mismatches: 8
Query Match: 34.85% Indels: 0
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x US-11-112-908-24 (1-150314)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||
Db 145200 CCACTGCATCTCCAGCCTAGTGACAGACCAAGACTCCGCTCTCAAAAAAAAAAAAAAGT 145141
   |||

Qy 21 Lys 21
   |||
Db 145140 AAA 145138

RESULT 33
US-11-117-187-212/c
; Sequence 212, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 212
; LENGTH: 163317
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-212
```





NAME/KEY: modified\_base  
LOCATION: (158929)..(158929)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (163528)..(163530)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (163550)..(163550)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (163765)..(163765)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (164000)..(164000)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (164047)..(164047)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (164084)..(164084)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (167233)..(167233)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (167236)..(167236)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170247)..(170247)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170249)..(170250)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170252)..(170253)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170259)..(170259)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170262)..(170263)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (170266)..(170266)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (174470)..(174470)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (174472)..(174472)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base

LOCATION: (174474)..(174474)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (179059)..(179060)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (179064)..(179064)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (179001)..(179001)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (197004)..(197005)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (197007)..(197007)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (200349)..(200349)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (200351)..(200351)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (212425)..(212426)  
OTHER INFORMATION: a, c, g, t, unknown or other  
FEATURE:  
US-11-121-086-31  
Alignment Scores:  
Pred. No.: 3.53e+05 Length: 218821  
Score: 46.00 Matches: 8  
Percent Similarity: 69.23% Conservative: 1  
Best Local Similarity: 61.54% Mismatches: 4  
Query Match: 34.85% Indels: 0  
DB: 7 Gaps: 0  
US-09-632-036F-6 (1-24) x US-11-121-086-31 (1-218821)  
QY 11 AspGlyThrGlnAlaGluLysCysSerLysProCys 23  
Db 217811 GAAGGTACACAAATGCCAGTAAGTGCATGAAGAGATGC 217773  
RESULT 36  
US-11-121-086-27/c  
Sequence 27, Application US/11121086  
Publication No. US20050266459A1  
GENERAL INFORMATION:  
APPLICANT: POULSEN, TIM S.  
APPLICANT: NIELSEN, KIRSTEN V.  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
FILE REFERENCE: 09138.6000-00000  
CURRENT APPLICATION NUMBER: US/11/121,086  
CURRENT FILING DATE: 2005-05-04  
PRIOR APPLICATION NUMBER: 60/567,570  
PRIOR FILING DATE: 2004-05-04  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 27  
LENGTH: 153142  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-121-086-27  
Alignment Scores:  
Pred. No.: 2.79e+05 Length: 153142

[illegible]

```

; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 194
; LENGTH: 92584
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-194

```

```

Alignment Scores:
Pred. No.: 1.85e+05 Length: 92584
Score: 45.00 Matches: 10
Percent Similarity: 47.83% Conservative: 1
Best Local Similarity: 43.48% Mismatches: 12
Query Match: 34.09% Indels: 0
DB: 7 Gaps: 0

```

US-09-632-036F-6 (1-24) x US-11-117-187-194 (1-92584)

```

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 30841 CTTTTCACGACCAACATTGACCGCATAGCTCGCGAACTCAGAGAGGAGAAACACAGT 30782
Qy 21 LysProCys 23
Db 30781 GAACCTTGT 30773

```

Search completed: December 11, 2005, 03:20:19  
Job time : 393.455 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 22:26:14 ; Search time 1957.09 Seconds  
(without alignments)  
573.755 Million cell updates/sec

Title: US-09-632-036F-6  
Perfect score: 132  
Sequence: 1 PLHNQEVTAEDGTQRAEKCCKPCA 24

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09632036/runat\_02122005\_104141\_25823/app\_query.fasta\_1.398  
-DB-EST\_QFMT-fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCPU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09632036 @CGN 1.1 5432 @runat\_02122005\_104141\_25823 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_hic.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_est7.\*  
9: gb\_gest1.\*  
10: gb\_gest2.\*  
11: gb\_gest3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	97.0	333	7	CN409749
2	128	97.0	356	5	BX479259 DKFZ0686M
3	128	97.0	376	7	CN409731 170005321
4	128	97.0	497	5	BX479184 DKFZ0686G
5	128	97.0	566	7	CN409736 170005321
6	128	97.0	582	3	BP313449 BP313449
7	128	97.0	583	3	BP315895 BP315895

8	128	97.0	589	7	CN409744
9	128	97.0	618	7	CV571007
c 10	128	97.0	670	1	AI906012
11	128	97.0	808	6	CA489534
12	128	97.0	894	2	BE746725
13	128	97.0	1005	5	BX402419
14	128	97.0	2004	4	CR592336
15	128	97.0	3444	11	DQ047381
16	128	97.0	3695	11	DQ047380
17	121	91.7	748	8	DN511011
18	117	88.6	614	1	AW370693
19	116	87.9	820	8	CV884121
c 20	115	87.1	455	2	BG991986
21	114	86.4	439	10	CL413651
22	114	86.4	539	6	CA544382
23	114	86.4	682	7	CF902067
c 24	114	86.4	854	11	CR119637
25	114	86.4	4323	4	AK031099
26	114	86.4	4463	4	AK083669
27	110	83.3	795	6	CA328613
28	109	82.6	964	5	BUL50809
29	108	81.8	614	7	CR767843
30	92	69.7	474	1	AL701569
31	85	64.4	490	2	BF995059
32	85	64.4	490	2	BF995204
c 33	85	64.4	552	7	CV349344
34	82	62.1	1016	3	BM562913
c 35	75	56.8	115	1	AI908047
c 36	63	47.7	286	2	BG954969
37	63	47.7	709	5	BX336780
c 38	62	47.0	708	10	CWS38167
39	62	47.0	1004	11	CNS060UA
40	61.5	46.6	552	9	BH011016
41	61	46.2	636	9	BZ205610
42	60	45.5	526	6	CA286377
c 43	60	45.5	601	6	CA140633
44	60	45.5	641	6	CF576665
c 45	60	45.5	643	5	BQ829433
c 46	60	45.5	668	10	CZ675091
c 47	60	45.5	829	6	CA070410
48	60	45.5	1209	6	CD048799
49	59	44.7	743	7	CO800753
c 50	59	44.7	868	5	BX391706
c 51	58	43.9	348	9	BZ755010
c 52	58	43.9	388	10	CG743040
c 53	58	43.9	804	1	AI635339
c 54	58	43.9	856	6	CF553201
55	58	43.9	874	10	AG848401
56	57	43.2	149	5	BQ965994
c 57	57	43.2	405	2	BB765881
c 58	57	43.2	427	2	BB675006
c 59	57	43.2	427	2	BB676604
c 60	57	43.2	535	10	CZ412019
61	57	43.2	751	2	BG825937
62	57	43.2	755	5	BU253274
c 63	57	43.2	848	2	BF262437
c 64	57	43.2	910	3	BI601086
c 65	56	42.4	328	9	BZ868523
c 66	56	42.4	353	9	AQ035868
c 67	56	42.4	400	2	BB669924
68	56	42.4	651	3	BJ661324
c 69	56	42.4	805	7	CV662265
c 70	56	42.4	815	3	BI818940
71	56	42.4	844	9	BZ980494
c 72	56	42.4	953	10	CNS011EP
c 73	55.5	42.0	476	9	AQ694702
c 74	55.5	42.0	783	10	CL861864
c 75	55	41.7	356	5	BQ539210
c 76	55	41.7	434	2	BB677696
77	55	41.7	559	3	BQ869495
78	55	41.7	566	3	BJ037389
c 79	55	41.7	585	1	AI677624
c 80	55	41.7	587	1	AU035223

CN409744	170004243
CV571007	oe07g11.y
AI906012	RC-BT105
CA489534	AGENCOURT
BE746725	601579159
BX402419	BX402419
CR592336	full-leng
DQ047381	Pan trogl
DQ047380	Homo sapi
DN511011	HL0202181
AW370693	QV1-BT026
CV884121	UI-M-FCO-
BG991986	MR2-HT116
CL413651	RPC144 43
CA544382	C0648D06-
CF902067	A0335H03-
CR119637	Reverse s
AK031099	Mus muscu
AK083669	Mus muscu
CA328613	UI-M-PYO-
BUL50809	AGENCOURT
CR767843	DKFZp469C
AL701569	DKFZp686A
BF995059	MR2-GN015
BF995204	MR3-GN015
CV349344	MR2-HT116
BM562913	AGENCOURT
AI908047	QV-BT161-
BG954969	CM4-CT065
BX336780	BX336780
CWS38167	OP_Ba001
AL415232	T7_end of
BH011016	ep72h09.g
BZ205610	CH230-293
CA286377	SCGSD204
CA140633	SCJFRT205
CF576665	MCSA203C0
BQ829433	LL6in2234
CZ675091	OM_Ba023
CA070410	SCSGAD100
CD048799	AGENCOURT
CO800753	AGENCOURT
BX391706	BX391706
BZ755010	PUPBK25TB
CG743040	ZMWBRC019
AI635339	tz80h12.x
CF553201	AGENCOURT
AG848401	Oryza sat
BQ965994	QHB23K05
BB765881	BB765881
BB675006	BB675006
BB676604	BB676604
CZ412019	1007722 R
BG825937	602748060
BU253274	603747073
BF262437	HV_CEA000
BI601086	603249577
BZ868523	CH240_226
AQ035868	CIT-HSP-2
BB669924	BB669924
BJ661324	BB661324
CV662265	LCPE04EX0
BI818940	603037434
BZ980494	FUGI797D
AL45634	Anopheles
AQ694702	HS 5473 B
CL861864	OR_CBA009
BQ539210	MEST612-E
BB677696	BB677696
BQ869495	QGD6F13.y
BJ037389	BJ037389
AI677624	605057H06
AU035223	AU035223

```

81 55 41.7 609 3 BI839047
82 55 41.7 658 7 CF950469
C 83 55 41.7 663 5 BY752074
C 84 55 41.7 664 3 BM339638
85 55 41.7 668 3 BJ666137
86 55 41.7 708 9 BZ415417
87 55 41.7 753 10 CG102034
88 55 41.7 830 7 CO960339
C 89 55 41.7 836 9 BZ980491
C 90 55 41.7 840 8 DR785590
91 55 41.7 864 2 BG320282
92 55 41.7 913 11 CNS033KP
93 55 41.7 1496 4 BC075711
94 54.5 41.3 435 2 BI051584
95 54.5 41.3 588 6 CD909087
96 54.5 41.3 616 6 CD925860
C 97 54.5 41.3 619 2 BE162709
C 98 54.5 41.3 650 10 AG019503
C 99 54.5 41.3 780 9 CC143331
C 100 54.5 41.3 956 2 BG033293

```

## ALIGNMENTS

```

RESULT 1
CN409749 333 bp mRNA linear EST 16-MAY-2004
LOCUS 17000532194973 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN409749
VERSION CN409749.1 GI:47396873
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 333)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 333 Std Error: 0.00.

```

## FEATURES

```

source
Location/Qualifiers
1..333
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/notes="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hes cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4.85e-10 Length: 333
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 7 Gaps: 0

```

```

US-09-632-036F-6 (1-24) x CN409749 (1-333)
QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 69 CCCCTGCAACCAAGAGGTGACAGAGGATGGACACACGCGGTGTGAGAAGTGCAGC 128
QY 21 LysProCysAla 24
Db 129 AAGCCCTGTGCC 140

```

## RESULT 2

```

BX479259 356 bp mRNA linear EST 04-SEP-2003
DKFZp686M18209 r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DKFZp686M18209 5', mRNA sequence.

```

```

ACCESSION BX479259
VERSION BX479259
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 356)
AUTHORS Blocker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Blocker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp686M18209) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

## FEATURES

```

source
Location/Qualifiers
1..356
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M18209"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 5.25e-10 Length: 356
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservative: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 5 Gaps: 0
US-09-632-036F-6 (1-24) x BX479259 (1-356)

```

```

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db 172 CCCCTGCAACCAAGAGGTGACAGAGGATGGACACACGCGGTGTGAGAAGTGCAGC 231
QY 21 LysProCysAla 24
Db 232 AAGCCCTGTGCC 243

```

RESULT 3  
 LOCUS CN409731 376 bp mRNA linear EST 16-MAY-2004  
 DEFINITION 17000532198194 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN409731  
 VERSION CN409731.1 GI:47396855  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 376)  
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 TITLE Transcriptome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
 PUBMED 15146197  
 COMMENT Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: tbrandenberger@geron.com  
 Insert Length: 376 Std Error: 0.00.  
 FEATURES  
 source Location/Qualifiers  
 1..376  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue types="embryonic stem cells, embryoid bodies  
 derived from H1, H7 and H9 cells"  
 /clone\_lib="GRN\_EB"  
 /notes="oligo dt primed, full-length enriched cDNA library  
 from embryoid body outgrowths derived from HES cell lines  
 H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
 conditions."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.6e-10 Length: 376  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-632-036F-6 (1-24) x CN409731 (1-376)  
 QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 |||||  
 Db 188 CCCCTGCACCAACAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGCAGC 247  
 |||||  
 QY 21 LysProCysAla 24  
 |||||  
 Db 248 AGCCCTGTGCC 259  
 |||||

RESULT 4  
 LOCUS BX479184 497 bp mRNA linear EST 04-SEP-2003  
 DEFINITION DKFZp686G01209.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone  
 ACCESSION DKFZp686G01209.5', mRNA sequence.  
 VERSION BX479184  
 KEYWORDS EST.  
 SOURCE BX479184.1 GI:31914729  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 497)  
 AUTHORS Bloecker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
 Pobo, G., Han, M. and Wiemann, S.  
 TITLE EST (Bloecker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., et al.)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: MIPS  
 MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 No sl sequence available.  
 This clone (DKFZp686G01209) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
 FEATURES  
 source Location/Qualifiers  
 1..497  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686G01209"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="686 (synonym: hlcc3)"  
 /note="Vector: pTriplex2; Site\_1: SfIIA; Site\_2: SfiIB;  
 cDNA-collection"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.76e-10 Length: 497  
 Score: 128.00 Matches: 23  
 Percent Similarity: 95.83% Conservative: 0  
 Best Local Similarity: 95.83% Mismatches: 1  
 Query Match: 96.97% Indels: 0  
 DB: 5 Gaps: 0  
 US-09-632-036F-6 (1-24) x BX479184 (1-497)  
 QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 |||||  
 Db 171 CCCCTGCACCAACAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGCAGC 230  
 |||||  
 QY 21 LysProCysAla 24  
 |||||  
 Db 231 AAGCCCTGTGCC 242  
 |||||

RESULT 5  
 LOCUS CN409736 566 bp mRNA linear EST 16-MAY-2004  
 DEFINITION 17000532198206 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN409736  
 VERSION CN409736.1 GI:47396860  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 566)  
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 TITLE Transcriptome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
 PUBMED 15146197  
 COMMENT Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 US-09-632-036F-6 (1-24) x BX479184 (1-497)  
 QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 |||||  
 Db 171 CCCCTGCACCAACAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGCAGC 230  
 |||||  
 QY 21 LysProCysAla 24  
 |||||  
 Db 231 AAGCCCTGTGCC 242  
 |||||

Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 566 Std Error: 0.00.

#### FEATURES

source  
1. .566  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN EB"  
/notes="Oligo dT primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

#### ORIGIN

##### Alignment Scores:

Pred. No.: 9,04e-10 Length: 566  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x CN409736 (1-566)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAapGlyThrGlnArgAlaGluLysCysSer 20  
DB 412 CCCCTGCACCAACAGAGGTGACAGCAGAGGTGGAACACACGCGGTGTGAGAAGTGCAGC 471  
QY 21 LysProCysAla 24  
DB 472 AAGCCCTGTGCC 483

#### RESULT 6

LOCUS BP313449 582 bp mRNA linear EST 17-SEP-2004  
DEFINITION BP313449 Sugano cDNA library, mammary gland OCUB-F Homo sapiens  
CDNA clone OFR03249, mRNA sequence.

ACCESSION BP313449  
VERSION BP313449.1 GI:52242424  
KEYWORDS EST.

SOURCE Homo sapiens (human)

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 582)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

15342556

CONTACT: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. .582

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="OFR03249"

/tissue\_type="mammary gland"

/cell\_line="OCUB-F"

/clone\_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

#### ORIGIN

##### Alignment Scores:

Pred. No.: 9,34e-10 Length: 582  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x BP313449 (1-582)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAapGlyThrGlnArgAlaGluLysCysSer 20  
DB 363 CCCCTGCACCAACAGAGGTGACAGCAGAGGTGGAACACACGCGGTGTGAGAAGTGCAGC 422  
QY 21 LysProCysAla 24  
DB 423 AAGCCCTGTGCC 434

#### RESULT 7

LOCUS BP315895 583 bp mRNA linear EST 17-SEP-2004  
DEFINITION BP315895 Sugano cDNA library, mammary gland OCUB-F Homo sapiens  
CDNA clone OFR09650, mRNA sequence.

ACCESSION BP315895  
VERSION BP315895.1 GI:52244870  
KEYWORDS EST.

SOURCE Homo sapiens (human)

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 583)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

15342556

CONTACT: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. .583

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="OFR09650"

/tissue\_type="mammary gland"

/cell\_line="OCUB-F"

/clone\_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

#### ORIGIN

##### Alignment Scores:

Pred. No.: 9,36e-10 Length: 583  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x BP315895 (1-583)

QY 1 ProLeuHisAenGlnGluValThrAlaGluAapGlyThrGlnArgAlaGluLysCysSer 20  
DB 415 CCCCTGCACCAACAGAGGTGACAGCAGAGGTGGAACACACGCGGTGTGAGAAGTGCAGC 474  
QY 21 LysProCysAla 24  
DB 475 AAGCCCTGTGCC 486



```

RESULT 8
LOCUS      CN409744
DEFINITION 589 bp mRNA linear EST 16-MAY-2004
ACCESSION 17000424385945 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
VERSION    CN409744
KEYWORDS   CN409744.1 GI:47396868
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 589)
AUTHORS    Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
            Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
            Lebkowski, J. and Stanton, L.W.
TITLE      Transcriptional characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
PUBMED     Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
FEATURES   source
            1..589
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, cell lines H1, H7, and
            H9"
            /clone_lib="GRN_ES"
            /note="oligo dt primed, full-length enriched cDNA library
            from undifferentiated hES cell lines H1 (p32), H7 (p29),
            and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Alignment Scores:
Pred. No.:      9,48e-10      Length:      589
Score:          128.00      Matches:    23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%      Mismatches: 1
Query Match:    96.97%      Indels:    0
DB:             7           Gaps:       0

US-09-632-036F-6 (1-24) x CN409744 (1-589)

Qy      1  ProLeuHisAnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      453  CCCCTGCACACCAAGAGGTGACAGCAGAGAGATGGAACACACAGCGGTGTGAGAGTGCAGC 512

Qy      21  LysProCysAla 24
Db      513  AAGCCCTGTGCC 524

RESULT 9
CV571007
LOCUS      CV571007
DEFINITION 618 bp mRNA linear EST 22-OCT-2004
ACCESSION CV571007
KEYWORDS   sapiens cDNA clone oe07g11 5', mRNA sequence.
SOURCE     CV571007
ORGANISM   Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 618)

```

```

AUTHORS    Rabinowitz, Y., Dong, L. and Wistow, G.
TITLE      Expressed sequence tag analysis of human keratoconus cornea
JOURNAL    Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 07 row: g column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES   source
            1..618
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="oe07g11"
            /tissue_type="Cornea"
            /dev_stage="Adult"
            /lab_host="EMDH10B"
            /clone_lib="Human keratoconus cornea, unamplified,
            (od/oe)"
            /note="Organ: Eye; Vector: pCMVSPORT6; Approximately 40ug
            total RNA was extracted from 7 adult human keratoconus
            corneas. A directionally cloned cDNA library in the
            pSPORT1 vector (Invitrogen) was constructed at Bioserve
            Biotechnology (Laurel MD) essentially following the
            protocols of the SuperScript Plasmid System full details
            of which are contained in the manufacturer's instruction
            manual (http://www.lifetech.com/). First strand synthesis
            was carried out using a Not I primer-adaptor
            [5'-pCAGTAGTTCTAGATCGAGCGGCCGCT(15-3')]. cDNA was
            cloned in Not I/Sal I sites. EST analysis was performed on
            the unamplified library at the NIH Intramural Sequencing
            Center (NISC)."
ORIGIN
Alignment Scores:
Pred. No.:      1e-09      Length:      618
Score:          128.00      Matches:    23
Percent Similarity: 95.83%      Conservative: 0
Best Local Similarity: 95.83%      Mismatches: 1
Query Match:    96.97%      Indels:    0
DB:             7           Gaps:       0

US-09-632-036F-6 (1-24) x CV571007 (1-618)

Qy      1  ProLeuHisAnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
Db      25  CCCTGCACACCAAGAGGTGACAGCAGAGATGGAACACACAGCGGTGTGAGAGTGCAGC 84

Qy      21  LysProCysAla 24
Db      85  AAGCCCTGTGCC 96

RESULT 10
AI906012/c
LOCUS      AI906012/c
DEFINITION RC-BT105-110399-035 BT105 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI906012
VERSION    AI906012.1 GI:6496399
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 670)
AUTHORS    Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,

```

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL  
PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED  
10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?tl=RC&t2=RC-BT105-035.html  
&t3=il0399&t4=1)  
Seq primer: puc 18 forward.

## FEATURES

source

Location/Qualifiers  
1..670  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stage="Adult"  
/clone\_lib="BT105"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORFEST PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Alignment Scores:

Pred. No.: 1..1e-09 Length: 670  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 1 Gaps: 0

US-09-632-036F-6 (1-24) x AI906012 (1-670)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysser 20  
|||||  
Db 157 CCCTCGCAACCAAGAGGTGACAGCAGAGATGGAACACACAGCGGTGAGAGTGCAGC 98  
|||||  
QY 21 LysProCySala 24  
|||||  
Db 97 AAGCCCTGTGCC 86  
|||||

## RESULT 11

CA489534

LOCUS

DEFINITION AGENCOURT\_10810594 MAPcL Homo sapiens cDNA clone IMAGE:6722008 5',  
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 808)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM14284 row: a column: 16  
High quality sequence stop: 582.

## FEATURES

source

Location/Qualifiers  
1..808  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6722008"  
/cell\_lines="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,  
hTERT-HME1, LNCap"  
/lab\_host="EMDH10B"  
/clone\_lib="MAPcL"  
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;  
Subtracted with brain, liver, lung, kidney and muscle.  
Directionally cloned. Priming method: oligo-dT. Average  
insert size: 1800 bp. Library amplification: 26,000 fold.  
Kristi A. Eglund, James J. Vincent, Robert Strausberg,  
Bungkook Lee & Ira Pastan: Discovery of new breast  
cancer genes encoding membrane and secreted proteins.  
Manuscript submitted."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.37e-09 Length: 808  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CA489534 (1-808)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysser 20  
|||||  
Db 552 CCCTCGCAACCAAGAGGTGACAGCAGAGATGGAACACACAGCGGTGAGAGTGCAGC 611  
|||||

QY 21 LysProCySala 24

|||||

Db 612 AAGCCCTGTGCC 623

## RESULT 12

BE746725

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 894)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM758 row: e column: 16

High quality sequence start: 3  
High quality sequence stop: 762.

## FEATURES

source  
1..894  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:3927927"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_9"  
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Alignment Scores:  
Pred. No.: 1 55e-09 Length: 894  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x BE746725 (1-894)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
Db 260 CCCCTGCACAAACAGAGGTGACAGCAGAGGTGGAACACAGCGGTGTGAGAGTGCAGC 319  
QY 21 LysProCysAla 24  
Db 320 AAGCCCTGTGCC 331

## RESULT 13

BX402419  
LOCUS BX402419 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CSODI032YB05 5-PRIME, mRNA sequence.

ACCESSION BX402419

VERSION BX402419.2 GI:46922471

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1005)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 13, 2003 this sequence version replaced gi:30632074.

## COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8568.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?e=CSIAI008ZH0JQP1&c=8568.f.

## FEATURES

source

1..1005

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="CSODI032YB05"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Alignment Scores:  
Pred. No.: 1 77e-09 Length: 1005  
Score: 128.00 Matches: 23  
Percent Similarity: 95.83% Conservative: 0  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 96.97% Indels: 0  
DB: 5 Gaps: 0

US-09-632-036F-6 (1-24) x BX402419 (1-1005)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20

Db 435 CCCCTGCACAAACAGAGGTGACAGCAGAGGTGGAACACAGCGGTGTGAGAGTGCAGC 494

QY 21 LysProCysAla 24

Db 495 AAGCCCTGTGCC 506

## RESULT 14

CR592336

LOCUS CR592336

DEFINITION full-length cDNA clone CSODI032YB05 of Placenta Cot 25-normalized

of Homo sapiens (human).

ACCESSION CR592336

VERSION CR592336.1 GI:50473143

KEYWORDS HTC; CNSLT\_cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2004)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue

2 (bases 1 to 2004)

Genoscope.

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

FEATURES

source

1..2004

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSODI032YB05"

/tissue\_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT\_6"

## ORIGIN

Alignment Scores:

Pred. No.: 3 99e-09 Length: 2004

Score: 128.00 Matches: 23

Percent Similarity: 95.83% Conservative: 0

```

Best Local Similarity: 95.83%      . Mismatches: 1
Query Match: 96.97%               Indels: 0
DB: 4                             Gaps: 0

US-09-632-036F-6 (1-24) x CR592336 (1-2004)

QY 1 ProLeuHisAsnGlnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 435 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC 494

QY 21 LysProCysAla 24
   |||||
Db 495 AAGCCCTGTGCC 506

RESULT 15
DQ047381
LOCUS DQ047381 3444 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes ERBB2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ047381.1 GI:66900580
VERSION DQ047381
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 3444)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 3444)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
1..3444
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3444
/gene="ERBB2"
/locus_tag="HC15304"

ORIGIN
Alignment Scores:
Pred. No.: 7.53e-09 Length: 3444
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x DQ047381 (1-3444)

QY 1 ProLeuHisAsnGlnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 873 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC 932

QY 21 LysProCysAla 24
   |||||
Db 933 AAGCCCTGTGCC 944

RESULT 17
DQ047381
LOCUS DQ047381 3444 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes ERBB2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ047381.1 GI:66900579
VERSION DQ047381
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3695)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 3695)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
1..3695
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3695
/gene="ERBB2"
/locus_tag="HC15304"

ORIGIN
Alignment Scores:
Pred. No.: 8.17e-09 Length: 3695
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x DQ047380 (1-3695)

QY 1 ProLeuHisAsnGlnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 873 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC 932

QY 21 LysProCysAla 24
   |||||
Db 933 AAGCCCTGTGCC 944

RESULT 17
DQ047380
LOCUS DQ047380 3695 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens ERBB2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ047380
VERSION DQ047380.1 GI:66900579
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3695)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 3695)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
1..3695
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3695
/gene="ERBB2"
/locus_tag="HC15304"

ORIGIN
Alignment Scores:
Pred. No.: 8.17e-09 Length: 3695
Score: 128.00 Matches: 23
Percent Similarity: 95.83% Conservatives: 0
Best Local Similarity: 95.83% Mismatches: 1
Query Match: 96.97% Indels: 0
DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x DQ047380 (1-3695)

QY 1 ProLeuHisAsnGlnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20
   |||||
Db 873 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGAACACACACGCGGTGTGAGAAGTGCAGC 932

QY 21 LysProCysAla 24
   |||||
Db 933 AAGCCCTGTGCC 944

RESULT 17
DQ047380
LOCUS DQ047380 3695 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens ERBB2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ047380
VERSION DQ047380.1 GI:60721201
KEYWORDS EST.
SOURCE Equus caballus (horse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 748)

```



/tissue\_type="whole brain"  
 /dev\_stage="embryo 12.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH BMAP FC0"  
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured mRNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was size selected according to mRNA size fraction,  
 ligated with EcoR I adaptor, digested with Not I, and then  
 cloned directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA  
 tail, is TGAGAGAGCC. This library was created for the  
 University of Iowa Mouse Brain Molecular Anatomy Project  
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous  
 System', supported by National Institutes of Mental Health  
 (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,11e-07 Length: 820  
 Score: 116.00 Matches: 21  
 Percent Similarity: 91.67% Conservative: 1  
 Best Local Similarity: 87.50% Mismatches: 2  
 Query Match: 87.88% Indels: 0  
 DB: 8 Gaps: 0

US-09-632-036F-6 (1-24) x CV884121 (1-820)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 ||| :|||  
 Db 30 CCCCGACACACACAGAGGTGCAGCTGAGGCGGACACACACGCGTGTGAGAAATGCAGC 89  
 ||| :|||  
 QY 21 LysProCysAla 24  
 ||| :|||  
 Db 90 AAGCCCTGTGCT 101  
 ||| :|||

RESULT 20  
 BG991986/c  
 LOCUS  
 DEFINITION MR2-HT1163--310101-016-f04 HT1163 Homo sapiens cDNA, mRNA sequence. EST 13-JUN-2001  
 ACCESSION BG991986  
 VERSION BG991986.1 GI:14396056  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 455)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## JOURNAL

## PUBMED

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?c1=MR2&t2=MR2-HT1163-

310101-016-f04&r3=2001-01-31&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 454.

## FEATURES

## source

1..455  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT1163"  
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Alignment Scores:  
 Pred. No.: 8.02e-08 Length: 455  
 Score: 115.00 Matches: 23  
 Percent Similarity: 92.00% Conservative: 0  
 Best Local Similarity: 92.00% Mismatches: 1  
 Query Match: 87.12% Indels: 1  
 DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x BG991986 (1-455)

QY 1 ProLeuHisAsn-GlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSe 20  
 ||| :|||  
 Db 126 CCCCTGTCACACACAGAGGTGCACAGCAGAGGTGTAACACACAGCGGTGTGAGAGTGCAG 67  
 ||| :|||

QY 20 rLysProCysAla 24  
 ||| :|||

Db 66 CAAGCCCTGTGCC 54  
 ||| :|||

## RESULT 21

## LOCUS

CL413651 419 bp DNA linear GSS 19-AUG-2004  
 DEFINITION RPCI44\_433J16.r RPCI-44 Sub scrofa genomic clone RPCI44\_433J16,  
 genomic survey sequence.

## ACCESSION

## VERSION

CL413651.1 GI:51447711

## KEYWORDS

## SOURCE

## ORGANISM

Sus scrofa (pig)  
 Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
 Sus.

1 (bases 1 to 419)

## REFERENCE

## AUTHORS

Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,  
 Beaver,J.E. and Schook,L.B.  
 TITLE Piggy-BACing the Human Genome: Constructing a Porcine Physical Map  
 Through Comparative Genomics

## JOURNAL

## COMMENT

Unpublished (2004)  
 Contact: Lawrence B. Schook  
 Department of Animal Sciences  
 University of Illinois at Urbana Champaign  
 1201 W. Gregory Dr., Urbana, IL 61801, USA  
 Tel: 217 265 5326  
 Fax: 217 244 5617

Email: schook@uiuc.edu

Clones are derived from the porcine BAC library RPCI-44  
 (http://www.bacpac.chori.org/porcine242.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@chori.org).  
 Clones may be purchased from BACPAC Resources  
 (http://BACPACorders.chori.org). This work was undertaken as part  
 of the International Swine Genome Sequencing Consortium by  
 University of Illinois at Urbana Champaign, USA with funds provided  
 by grant No. AG2002-34480-11828 from USDA-CSREES and  
 AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing  
 Initiative)

Plate: 433 row: J column: 16  
Seq primer: SP6  
Class: BAC ends.

## FEATURES

Location/Qualifiers  
source  
1..419  
/organism="Sus scrofa"  
/mol\_type="genomic DNA"  
/strain="Four pigs (breed: 37.5% Yorks Landrace and 25% Meishan)"  
/db\_xref="taxon:9823"  
/clone="RPC144\_433J16"  
/sex="male"  
/cell\_type="blood"  
/notes="Vector: pTARBAC2; Site 1: EcoRI; Site 2: EcoRI; porcine male BAC library produced by Pieter de Jong"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,05e-07 Length: 419  
Score: 114.00 Matches: 21  
Percent Similarity: 91.67% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 2  
Query Match: 86.36% Indels: 0  
DB: 10 Gaps: 0

US-09-632-036F-6 (1-24) x CL413651 (1-419)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
||| :|||  
Db 117 CCCCCGAAACCAACGAGGTGACAGCTGAGGACGGAACGCGGTGTGTGAGAAATGCAGC 176  
QY 21 LysProCysAla 24  
||| :|||  
Db 177 AAGCCCTGTGCC 188

## RESULT 22

CA544382 539 bp mRNA linear EST 19-NOV-2002  
LOCUS C0648D06-5N NIA Mouse Trophoblast Stem Cell cDNA Library (Long) Mus  
DEFINITION musculus cDNA clone NIA:C0648D06 IMAGE:30024809 5', mRNA sequence.

CA544382 GI:25087132

## ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE  
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Luo, A.,  
Tanaka, T., Kunath, T., Rossant, J. and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Trophoblast Stem Cell cDNA Library  
(Long)

Unpublished (2001)

## JOURNAL

COMMENT

Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgaun.grc.nia.nih.gov

Plate: C0648 row: D column: 06

Seq primer: M13 Reverse

High quality sequence stop: 539

POLYA=No.

Location/Qualifiers

1..539

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="B5/EGFP transgenic ICR mice"

/db\_xref="niaEST:C0648D06-5N"

/db\_xref="taxon:10090"

/clone="NIA:C0648D06 IMAGE:30024809"

/tissue\_type="Trophoblast stem cell"

/dev\_stage="3.5-dpc"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Trophoblast Stem Cell cDNA Library  
(Long)"  
/notes="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://lgaun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). (PMID: 11544199)). Total RNAs were  
obtained from Dr. Janet Rossant and Tilo Kunath (Samuel  
Lunenfeld Research Institute, Canada). Double-stranded  
cDNAs were synthesized with an Oligo(dT) primer  
(Invitrogen):  
5'-pGACTAGTCTAGATCGAGCGCGCCCTTTT-3' from  
4 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to lone-linker LL-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Taq polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.6 kb. The library was constructed  
by Yulan Piao (NIA)."

## ORIGIN

Alignment Scores:  
Pred. No.: 1,41e-07 Length: 539  
Score: 114.00 Matches: 21  
Percent Similarity: 91.67% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 2  
Query Match: 86.36% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CA544382 (1-539)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
||| :|||  
Db 31 CCCCCGAAACCAACGAGGTGACAGCTGAGGACGGAACGCGGTGTGTGAGAAATGCAGC 90

QY 21 LysProCysAla 24  
||| :|||

Db 91 AAGCCCTGTGCT 102

## RESULT 23

CF902067

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 682)

Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

11544199

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

```
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0335 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 682
POLYA=No.

FEATURES
    source
        1..682
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="B5/EGFP transgenic ICR mice"
            /db_xref="niaEST:A0335H03-5"
            /db_xref="taxon:10090"
            /clone="NIA:A0335H03 IMAGE:30736598"
            /dev_stage="3.5-dpc"
            /lab_host="DH10B"
            /clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library (Long 1)"
            /notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Janet Rossant and Tilo Kunath (Samuel Lunenfeld Research Institute, Canada). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen): 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3' from 4 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.6 kb. The library was constructed by Yulan Piao."
```

US-09-632-036F-6 (1-24) x CF902067 (1-682)

QY 1 ProLeuHisAsnGlnuValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
||| :|||  
Db 54 CCCCGAACAACCAAGAGGTACAGCGGAGGACGACAGCGGTGTGAGAATGCAGC 113

QY 21 LysProCysAla 24  
|||  
Db 114 AAACCTGTGCT 125

RESULT 24  
CR119637/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CR119637 854 bp DNA linear GSS 06-JUL-2004  
Reverse strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP345a06, genomic survey sequence.  
CR119637  
CR119637.1 GI:49867070  
GSS: genome survey sequence; M1CER.  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
1 (bases 1 to 854)  
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.  
Direct Submission  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
Location/Qualifiers  
1..854  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP345a06"  
/clone\_lib="MHPp"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.42e-07 Length: 854  
Score: 114.00 Matches: 21  
Percent Similarity: 91.67% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 2  
Query Match: 86.36% Indels: 0  
DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x CR119637 (1-854)

QY 1 ProLeuHisAsnGlnuValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
||| :|||  
Db 408 CCCCGAACAACCAAGAGGTACAGCGGAGGACGACAGCGGTGTGAGAATGCAGC 349

QY 21 LysProCysAla 24  
|||  
Db 348 AAGCCTGTGCT 337

RESULT 25  
AK031099  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK031099 4323 bp mRNA linear HTC 03-APR-2004  
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length  
enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic  
leukemia viral oncogene homolog 2, neuro/glioblastoma derived  
oncogene homolog (avian), full insert sequence.  
AK031099  
AK031099.1 GI:26082143  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

CR119637/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CR119637 854 bp DNA linear GSS 06-JUL-2004  
Reverse strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP345a06, genomic survey sequence.  
CR119637  
CR119637.1 GI:49867070  
GSS: genome survey sequence; M1CER.  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;





Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/.

## FEATURES

source

Location/Qualifiers  
1. .4463

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:D030063B12"  
/db\_xref="taxon:10090"  
/clone="D030063B12"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="9 days embryo"

## misc\_feature

1. .4463  
/note="v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGD|MG1:95410, GB|U71126, evidence: BLASTN, 99%, match=449)"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.68e-06 Length: 4463  
Score: 114.00 Matches: 21  
Percent Similarity: 91.67% Conservative: 1  
Best Local Similarity: 87.50% Mismatches: 2  
Query Match: 86.36% Indels: 0  
DB: 4 Gaps: 0

US-09-632-036F-6 (1-24) x AK083669 (1-4463)

QY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
|||:::|||||||||||||||||||||||||||||||||||||||||||||||||||

Db 996 CCCCAGACACACAGAGGTGACAGCTGAGGACGGAACACAGCGGTGTGAGAAATGCAC 1055  
|||||

QY 21 LysProCysAla 24  
|||||

Db 1056 AAGCCCTGTGCT 1067  
|||||

## RESULT 27

CA328613

LOCUS

DEFINITION UI-M-FYO-cda-e-16-0-UI.r1 NIH\_BMAP\_FYO Mus musculus cDNA clone  
IMAGE:6826841 5', mRNA sequence.

ACCESSION

CA328613

VERSION

CA328613.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

NIH-MGC

http://mgc.nci.nih.gov/.

Unpublished (1999)

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

## FEATURES

source

Location/Qualifiers  
1. .795

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:6826841"

/tissue\_type="whole brain"

/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab\_host="DH10B (T1 phage resistant)"

/clone\_lib="NIH\_BMAP\_FYO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is ACCGAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

## ORIGIN

Alignment Scores:  
Pred. No.: 9.55e-07 Length: 795  
Score: 110.00 Matches: 20  
Percent Similarity: 95.45% Conservative: 1  
Best Local Similarity: 90.91% Mismatches: 1  
Query Match: 83.33% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-6 (1-24) x CA328613 (1-795)

QY 3 HisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerIysPro 22  
:::|||||

Db 703 AACAAACCAAGAGGTGACAGCTGAGGACGGAACACACAGCGGTGTGAGAAATGCACGACCC 762  
|||||

QY 23 CysAla 24  
|||||

Db 763 TGTGCT 768  
|||||

## RESULT 28

BU150809

LOCUS

DEFINITION

AGENCY: 8118591 Lupeki dorsal root ganglion Homo sapiens cDNA

clone IMAGE:6180101 5', mRNA sequence.

ACCESSION

BU150809

VERSION

BU150809.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

NIH-MGC

http://mgc.nci.nih.gov/.

Unpublished (1999)

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
cDNA Library preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAMI3562 row: f column: 06  
 High quality sequence start: 27  
 High quality sequence stop: 529.

#### FEATURES

Location/Qualifiers  
 1..964  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6180101"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site\_1:  
 Note1: Site2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGTCCG-3' and  
 5'-GACTAGTCTTAGATCGGCGCGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.7 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1.72e-06 Length: 964  
 Score: 109.00 Matches: 20  
 Percent Similarity: 95.24% Conservative: 0  
 Best Local Similarity: 95.24% Mismatches: 1  
 Query Match: 82.58% Indels: 0  
 DB: 5 Gaps: 0

US-09-632-036F-6 (1-24) x BU150809 (1-964)

OY 4 AenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCys 23  
 Db 33 ACCCAAGAGTGACACAGAGGATGGACACACCGGTGTGAGAAGTGCACGACCGCTGT 92  
 OY 24 Ala 24  
 Db 93 GCC 95

RESULT 29  
 CR767843 614 bp mRNA linear EST 23-SEP-2004  
 DKFZp469C0110 r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone  
 DEFINITION  
 CR767843  
 CR767843.1 GI:52610368  
 EST.  
 SOURCE  
 Pongo pygmaeus (orangutan)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Pongo.  
 REFERENCE  
 1 (bases 1 to 614)  
 OTtenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,  
 Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and  
 Wiemann,S.  
 Pongo pygmaeus mRNA (Otenwaelder,B., Obermaier,B.,  
 Deutschenbaur,S., et al.)  
 Unpublished (2004)  
 CONTACT: MIPS  
 MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix  
 (Martineried/Germany) within the cDNA sequencing consortium of the

German Genome Project. This clone (DKFZp469C0110) is available at  
 the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in  
 Berlin, Germany. Please contact RZPD for ordering:  
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C0110  
 Further information about the clone and the sequencing project is  
 available at http://mips.gsf.de/projects/cdna/.

#### FEATURES

Location/Qualifiers  
 1..614  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp469C0110"  
 /tissue\_type="kidney"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="469 (synonym: pkid1)"  
 /note="Vector: pSPORT1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1.46e-06 Length: 614  
 Score: 108.00 Matches: 20  
 Percent Similarity: 95.24% Conservative: 0  
 Best Local Similarity: 95.24% Mismatches: 1  
 Query Match: 81.82% Indels: 0  
 DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x CR767843 (1-614)

OY 1 ProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 Db 552 CCCCTGTCACCAACAGAGGTGACAGCAGAGGACGACACACGCGGTGTGAGAAGTGCAGC 611  
 OY 21 Lys 21  
 Db 612 AAG 614

#### RESULT 30

AL701569 474 bp mRNA linear EST 04-SEP-2003  
 DKFZp686A19145 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
 DEFINITION  
 AL701569  
 AL701569.1 GI:19684925  
 EST.  
 KEYWORDS  
 SOURCE  
 Homo sapiens (human)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 474)  
 Bloeker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B. and  
 Wiemann,S.  
 EST (Bloeker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B. and  
 Wiemann,S.)  
 Unpublished (1999)  
 CONTACT: MIPS  
 MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.

No si sequence available.  
 This clone (DKFZp686A19145) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

#### FEATURES

Location/Qualifiers  
 1..474  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"



Alignment Scores:  
 Pred. No.: 0.00493 Length: 490  
 Score: 85.00 Matches: 17  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 80.95% Mismatches: 3  
 Query Match: 64.39% Indels: 0  
 DB: 2 Gaps: 0

US-09-632-036F-6 (1-24) x BF995204 (1-490)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 |||||  
 Db 73 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGCAACACAGAGGTGGAGGAGGAGAG 132  
 |||||

Qy 21 Lys 21  
 |||||  
 Db 133 AAA 135

RESULT 33  
 CV349344/c  
 LOCUS CV349344 552 bp mRNA linear EST 24-SEP-2004  
 DEFINITION MR2-HT1163-130201-013-f02 HT1163 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION CV349344  
 VERSION CV349344.1 GI:52672558  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 1 (bases 1 to 552)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. http://www.ludwig.org.br.

FEATURES  
 source  
 1..552  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT1163"  
 /notes="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00567 Length: 552  
 Score: 85.00 Matches: 18  
 Percent Similarity: 76.00% Conservative: 1  
 Best Local Similarity: 72.00% Mismatches: 5

Query Match: 64.39% Indels: 1  
 DB: 7 Gaps: 0

US-09-632-036F-6 (1-24) x CV349344 (1-552)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSer 20  
 |||||  
 Db 122 CCCCTGCACACCAAGAGGTGACAGCAGAGGATGGACACACCGCGTGTGAGAGTGCAGC 63  
 |||||

Qy 21 Lys-ProCysAla 24  
 |||||  
 Db 62 AACGCCCTGTGCC 50  
 |||||

RESULT 34  
 BM562913  
 LOCUS BM562913 1016 bp mRNA linear EST 20-FEB-2002  
 DEFINITION AGENCOURT\_6566679 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5736771  
 5', mRNA sequence.  
 ACCESSION BM562913  
 VERSION BM562913.1 GI:18809393  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 1 (bases 1 to 1016)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Straubeberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: ARCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12745 row: n column: 04  
 High quality sequence stop: 637.

FEATURES  
 source  
 1..1016  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5736771"  
 /tissue\_type="duodenal adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_88"  
 /note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.0346 Length: 1016  
 Score: 82.00 Matches: 19  
 Percent Similarity: 79.50% Conservative: 2  
 Best Local Similarity: 79.17% Mismatches: 1  
 Query Match: 62.12% Indels: 2  
 DB: 3 Gaps: 0

US-09-632-036F-6 (1-24) x BM562913 (1-1016)

Qy 1 ProLeuHisAenGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCys 20  
 |||||  
 Db 797 CCCCTGCACACCAAGAGGTGACCGAGAAGATGGAACACACACCGCGTGTGAAAAAGTGCA 856  
 |||||

Qy 20 erLysPro 22  
 |||||

Db	857	ACCAGCCC	864	
RESULT 35				
AI908047/c				
LOCUS	AI908047	115 bp	mRNA	linear EST 30-MAR-2000
DEFINITION	QV-BT161-220499-124	BT161	Homo sapiens cDNA,	mRNA sequence.
ACCESSION	AI908047			
VERSION	AI908047.1	GI:6498727		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.			
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (7),	3491-3496	(2000)
PUBMED	10737800			
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/seq/gethtml.pl?tl=QV&t2=QV-BT161-124.html&t3=220499&t4=1) Seq primer: puc 18 forward.			
FEATURES	source			
	1..115	Location/Qualifiers		
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/sex="female"			
	/dev_stage="Adult"			
	/clone_lib="BT161"			
	/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."			
ORIGIN				
Alignment Scores:				
Pred. No.:	0.0345	Length:	115	
Score:	75.00	Matches:	13	
Percent Similarity:	92.86%	Conservative:	0	
Best Local Similarity:	92.86%	Mismatches:	1	
Query Match:	56.82%	Indels:	0	
DB:	1	Gaps:	0	
US-09-632-036F-6 (1-24) x AI908047 (1-115)				
QY	11	AspGlyThrGlnArgAlaGluLysCysSerLysProCysAla	24	
Db	113	GATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCC	72	
RESULT 36				
BG954969				
LOCUS	BG954969	266 bp	mRNA	linear EST 12-JUN-2001
DEFINITION	CM4-CT0656-120201-854-f06	CT0656	Homo sapiens cDNA,	mRNA sequence.

ACCESSION	BG954969			
VERSION	BG954969.1	GI:14373140		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 266)			
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (7),	3491-3496	(2000)
PUBMED	10737800			
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-CT0656-120201-854-f06&t3=2001-02-12&t4=1) Seq primer: puc 18 forward High quality sequence start: 11 High quality sequence stop: 266.			
FEATURES	source			
	1..266	Location/Qualifiers		
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/dev_stage="Adult"			
	/clone_lib="CT0656"			
	/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."			
ORIGIN				
Alignment Scores:				
Pred. No.:	7.34	Length:	266	
Score:	63.00	Matches:	11	
Percent Similarity:	91.67%	Conservative:	0	
Best Local Similarity:	91.67%	Mismatches:	1	
Query Match:	47.73%	Indels:	0	
DB:	2	Gaps:	0	
US-09-632-036F-6 (1-24) x BG954969 (1-266)				
QY	13	ThrGlnArgAlaGluLysCysSerLysProCysAla	24	
Db	264	ACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCC	229	
RESULT 37				
BX336780				
LOCUS	BX336780	709 bp	mRNA	linear EST 08-APR-2004
DEFINITION	BX336780	Homo sapiens	PLACENTA COT 25-NORMALIZED	Homo sapiens cDNA
		clone	CSODI032YB05 5-PRIME,	mRNA sequence.
ACCESSION	BX336780			
VERSION	BX336780.2	GI:46282048		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 709)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 2, 2003 this sequence version replaced gi:30345544.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8568.f  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1032CA03QPl&c=8568.f.  
FEATURES  
source 1..709  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1032YB05"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Alignment Scores:  
Pred. No.: 23.2 Length: 709  
Score: 63.00 Matches: 12  
Percent Similarity: 68.18% Conservative: 3  
Best Local Similarity: 54.55% Mismatches: 7  
Query Match: 47.73% Indels: 0  
DB: 5 Gaps: 0  
US-09-632-036F-6 (1-24) x BX336780 (1-709)  
QY 2 LeuHisAsnGlnClnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLys 21  
Db 202 CTGTACACCCAGAGGTGCATCGATGATGTWACAYAGYGKTKGAGWATTGTAGYWG 261  
QY 22 ProCys 23  
Db 262 CYCTGT 267  
RESULT 38  
CW538167/c  
LOCUS  
DEFINITION OP\_Ba0018C19.x OP\_Ba Oryza punctata genomic clone OP\_Ba0018C19  
3', Genomic survey sequence.  
ACCESSION CW538167  
VERSION CW538167.1 GI:54012389  
KEYWORDS GSS.  
SOURCE Oryza punctata  
ORGANISM Oryza punctata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 708)  
AUTHORS SanMiguel, P., Westerman, R., Kim, H., Yu, Y., Wissotski, M., Yost, D.,  
Stum, D., Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C.,  
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.  
OMAP Project - Purdue University  
Unpublished (2004)  
Contact: Scott A. Jackson

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 709)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 2, 2003 this sequence version replaced gi:30345544.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8568.f  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1032CA03QPl&c=8568.f.  
FEATURES  
source 1..709  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1032YB05"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Alignment Scores:  
Pred. No.: 23.2 Length: 709  
Score: 63.00 Matches: 12  
Percent Similarity: 68.18% Conservative: 3  
Best Local Similarity: 54.55% Mismatches: 7  
Query Match: 47.73% Indels: 0  
DB: 5 Gaps: 0  
US-09-632-036F-6 (1-24) x BX336780 (1-709)  
QY 2 LeuHisAsnGlnClnValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLys 21  
Db 202 CTGTACACCCAGAGGTGCATCGATGATGTWACAYAGYGKTKGAGWATTGTAGYWG 261  
QY 22 ProCys 23  
Db 262 CYCTGT 267  
RESULT 38  
CW538167/c  
LOCUS  
DEFINITION OP\_Ba0018C19.x OP\_Ba Oryza punctata genomic clone OP\_Ba0018C19  
3', Genomic survey sequence.  
ACCESSION CW538167  
VERSION CW538167.1 GI:54012389  
KEYWORDS GSS.  
SOURCE Oryza punctata  
ORGANISM Oryza punctata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 708)  
AUTHORS SanMiguel, P., Westerman, R., Kim, H., Yu, Y., Wissotski, M., Yost, D.,  
Stum, D., Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C.,  
Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.  
OMAP Project - Purdue University  
Unpublished (2004)  
Contact: Scott A. Jackson

Jackson Laboratory  
Purdue University  
915 W. State St., West Lafayette, IN 47907, USA  
Tel: 7654963621  
Fax: 7654967255  
Email: sjackson@purdue.edu  
This sequence was derived  
from the raw sequence read by clipping with lacy version 1.19s.  
Bases 129-836 of the raw sequence (length 1377) were retained after  
clipping.  
PCR Primers  
FORWARD: TAA TAC GAC TCA CTA TAG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Insert Length: 161000 Std Error: 0.00  
Platform: 0018 row: C column: 19  
Seq primer: CAC TCA TTA GGC ACC CCA  
Class: BAC ends.  
Location/Qualifiers  
1..708  
/organism="Oryza punctata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4537"  
/clone="OP\_Ba0018C19"  
/tissue\_type="young leaves"  
/lab\_host="DH10B-T1 phage resistant"  
/clone\_lib="OP\_Ba"  
/note="Vector: pAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN  
Alignment Scores:  
Pred. No.: 33.3 Length: 708  
Score: 62.00 Matches: 10  
Percent Similarity: 80.95% Conservative: 7  
Best Local Similarity: 47.62% Mismatches: 4  
Query Match: 46.97% Indels: 0  
DB: 10 Gaps: 0  
US-09-632-036F-6 (1-24) x CW538167 (1-708)  
QY 4 AsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCys 23  
Db 304 GATGAGGAAGTCAGTAAGGAGGAGGTCACCGACATGCTGACTTCTGTCTCGACCTGT 245  
QY 24 Ala 24  
Db 244 GCA 242  
RESULT 39  
CNS06U0A  
LOCUS  
DEFINITION T7 end of clone AX0AA003H07 of library AX0AA from strain CBS 7064  
of Fichia farinosa, genomic survey sequence.  
ACCESSION AL415232  
VERSION AL415232.1 GI:12194198  
KEYWORDS GSS.  
SOURCE Pichia farinosa  
ORGANISM Pichia farinosa  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Pichia.  
REFERENCE 1 (bases 1 to 1004)  
AUTHORS Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,  
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Caesaregola, S.,  
DeMontigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,  
Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Winkler, P. and Weissbach, J.  
Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies  
FEBS Lett. 487 (1), 3-12 (2000)  
11152876  
JOURNAL 2 (bases 1 to 1004)  
PUBMED  
REFERENCE 2 (bases 1 to 1004)  
AUTHORS de Montigny, J., Spohnner, C., Souciet, J., Tekala, F., Dujon, B.,  
Winkler, P., Artiguenave, P. and Potier, S.

**TITLE** Genomic exploration of the hemiascomycetous yeasts: 15. Pichia sorbitophila  
**JOURNAL** FEBS Lett. 487 (1), 87-90 (2000)  
**PUBMED** 11152890  
**REFERENCE** 3 (bases 1 to 1004)  
**AUTHORS** Genoscope.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)  
**COMMENT** This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

**FEATURES** source  
 1..1004  
 /organism="Pichia farinosa"  
 /mol\_type="genomic DNA"  
 /strain="CBS 7064"  
 /db\_xref="taxon:4920"  
 /clone="AX0AA003H07"  
 /clone\_lib="AX0AA"  
 /note="end : T7"

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 50.2 Length: 1004  
 Score: 62.00 Matches: 12  
 Percent Similarity: 75.00% Conservative: 3  
 Best Local Similarity: 60.00% Mismatches: 5  
 Query Match: 46.97% Indels: 0  
 DB: 11 Gaps: 0

US-09-632-036F-6 (1-24) x CNS06UOA (1-1004)

**QY** 5 GlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLysProCysAla 24  
 :::::::::::::::::::::  
**Db** 903 AGATCAGTAAACGGCGCGCGAGACGCTGCTGAAAAGTGTGGCGAGCCAGTGGC 962  
 :::::::::::::::::::::

**RESULT 40**  
**LOCUS** BH011016 552 bp DNA linear GSS 04-MAY-2001  
**DEFINITION** ep72h09.g1 TO1000 Brassica oleracea genomic clone ep72h09 5', genomic survey sequence.  
**ACCESSION** BH011016  
**VERSION** BH011016.1 GI:13957211  
**KEYWORDS** GSS.  
**SOURCE** Brassica oleracea  
**ORGANISM** Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

**REFERENCE** 1 (bases 1 to 552)  
**AUTHORS** Katari, M., O'Shaughnessy, A., Palmer, L., Bahret, A., Baker, J., Ballija, V., Bell, M., Cunnius, D.M., King, L., Kirchoff, K., Kuit, K., Miller, B., Nascimento, L., Preston, R., Rodriguez, S., Santos, L., Shah, R., Vil, M.D., Zutavern, T., Bal, H., Dedhia, N. and McCombie, W.R.  
 Whole Genome Shotgun Reads from Brassica oleracea Unpublished (2001)  
 Contact: W. Richard McCombie  
 Lita Annenberg Hazen Genome Sequencing Center  
 Cold Spring Harbor Laboratory  
 PO Box 100, Cold Spring Harbor, NY 11724, USA  
 Tel: 516 367 8884  
 Fax: 516 367 8874  
 Email: mcombie@cshl.org  
 Plate: ep72 row: h column: 09

**Seq primer:** -21revUniv  
**Class:** shotgun  
**High quality sequence stop:** 552.  
**FEATURES** source  
 1..552  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone="ep72h09"  
 /clone\_lib="TO1000"  
 /note="Vector: M13 for .x reads, pZero-2 for .b and .g reads; Site 1: EcoRV; DNA prepared as whole genome shotgun library from young, green leaves. May contain some plasmid DNAs. DNA provided by Dr. Tom Osborn, University of Wisconsin-Madison, Department of Agronomy."

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 29.9 Length: 552  
 Score: 61.50 Matches: 13  
 Percent Similarity: 66.67% Conservative: 1  
 Best Local Similarity: 61.90% Mismatches: 6  
 Query Match: 46.53% Indels: 1  
 DB: 9 Gaps: 1

US-09-632-036F-6 (1-24) x BH011016 (1-552)

**QY** 2 LeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgAlaGluLysCysSerLys 21  
 :::::::::::::::::::::  
**Db** 66 ATCCATAACGTTTCTTAGTACAGTTGAGATGGGCTCAACACGACGACAAATGT---AAA 122  
 :::::::::::::::::::::

**QY** 22 Pro 22  
**Db** 123 CCT 125

Search completed: December 11, 2005, 01:00:04  
 Job time : 1969.09 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:34:18 ; Search time 125.455 Seconds  
(without alignments)  
70.046 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSCVLDLDDKCPAEQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	100.0	22	4	AAB68633
2	105	92.1	30	6	ABJ18655
3	105	92.1	30	6	ABP59259
4	105	92.1	42	3	AAAY84783
5	105	92.1	42	3	AAAY84784
6	105	92.1	69	2	AAW53132
7	105	92.1	149	6	ABJ18657
8	105	92.1	149	6	ABP59261
9	105	92.1	435	8	ADP10480
10	105	92.1	470	9	ADY30515
11	105	92.1	493	7	ADG17936
12	105	92.1	653	3	AAB21200
13	105	92.1	653	5	AAM51145
14	105	92.1	675	8	ADW87401
15	105	92.1	678	8	ADT50884
16	105	92.1	678	9	ADZ47795
17	105	92.1	680	9	ADY28115
18	105	92.1	680	7	ADG17945
19	105	92.1	696	9	ADW28504
20	105	92.1	712	3	AAB21204
21	105	92.1	712	5	AAM51149
22	105	92.1	715	7	ADG17944
23	105	92.1	717	9	ADZ47807
24	105	92.1	720	7	ADG17940

25	105	92.1	741	7	ADM29356
26	105	92.1	782	2	AAW19764
27	105	92.1	815	8	Adt50879
28	105	92.1	835	9	ADY30511
29	105	92.1	879	9	ADY30514
30	105	92.1	919	3	AAB21203
31	105	92.1	919	5	AAM51148
32	105	92.1	951	3	AAV44993
33	105	92.1	960	9	ADY30513
34	105	92.1	970	9	ADY30509
35	105	92.1	1042	7	ADM29352
36	105	92.1	1200	3	AAB21208
37	105	92.1	1217	8	ADT50876
38	105	92.1	1233	5	AU98923
39	105	92.1	1253	7	ADC35106
40	105	92.1	1255	2	AAW01111
41	105	92.1	1255	2	AAW92406
42	105	92.1	1255	3	AAW92406
43	105	92.1	1255	3	AAW92406
44	105	92.1	1255	3	AAW92406
45	105	92.1	1255	3	AAW92406
46	105	92.1	1255	4	AAW92406
47	105	92.1	1255	4	AAW92406
48	105	92.1	1255	4	AAW92406
49	105	92.1	1255	4	AAW92406
50	105	92.1	1255	5	AAU77114
51	105	92.1	1255	5	AAU77114
52	105	92.1	1255	5	AAU77114
53	105	92.1	1255	5	AAU77114
54	105	92.1	1255	5	AAU77114
55	105	92.1	1255	5	AAU77114
56	105	92.1	1255	5	AAU77114
57	105	92.1	1255	6	ABP74708
58	105	92.1	1255	6	ABP74708
59	105	92.1	1255	6	ABP74708
60	105	92.1	1255	6	ABP74708
61	105	92.1	1255	6	ABP74708
62	105	92.1	1255	6	ABP74708
63	105	92.1	1255	6	ABP74708
64	105	92.1	1255	6	ABP74708
65	105	92.1	1255	6	ABP74708
66	105	92.1	1255	6	ABP74708
67	105	92.1	1255	6	ABP74708
68	105	92.1	1255	6	ABP74708
69	105	92.1	1255	6	ABP74708
70	105	92.1	1255	6	ABP74708
71	105	92.1	1255	6	ABP74708
72	105	92.1	1255	6	ABP74708
73	105	92.1	1255	6	ABP74708
74	105	92.1	1255	6	ABP74708
75	105	92.1	1255	6	ABP74708
76	105	92.1	1255	6	ABP74708
77	105	92.1	1255	6	ABP74708
78	105	92.1	1255	6	ABP74708
79	105	92.1	1255	6	ABP74708
80	105	92.1	1255	6	ABP74708
81	105	92.1	1255	6	ABP74708
82	105	92.1	1255	6	ABP74708
83	105	92.1	1255	6	ABP74708
84	105	92.1	1255	6	ABP74708
85	105	92.1	1255	6	ABP74708
86	105	92.1	1255	6	ABP74708
87	105	92.1	1255	6	ABP74708
88	105	92.1	1255	6	ABP74708
89	105	92.1	1255	6	ABP74708
90	105	92.1	1255	6	ABP74708
91	105	92.1	1255	6	ABP74708
92	105	92.1	1255	6	ABP74708
93	105	92.1	1255	6	ABP74708
94	105	92.1	1255	6	ABP74708
95	105	92.1	1255	6	ABP74708
96	105	92.1	1255	6	ABP74708
97	105	92.1	1255	6	ABP74708

98 105 92.1 1414 9 ADY30510 Human spl  
99 105 92.1 1433 2 AAR39568 Sequence  
100 100 87.7 1255 8 ADQ75969 Rhesus mo

## ALIGNMENTS

```
RESULT 1
AAB68633
ID AAB68633 standard; peptide; 22 AA.
XX
AC AAB68633;
XX
DT 30-APR-2001 (first entry)
XX
DE HER-2 B cell epitope #12.
XX
KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;
KW ovarian; lung; prostate; colon.
XX
OS Homo sapiens.
XX
PN WO200108636-A2.
XX
PD 08-FEB-2001.
XX
PF 03-AUG-2000; 2000WO-US021222.
XX
PR 03-AUG-1999; 99US-0146869P.
XX
PA (OHIS ) UNIV OHIO STATE.
XX
PI Kaumaya PT, Stevens VC, Triozzi PL;
XX
DR WPI; 2001-182849/18.
XX
PT Compositions comprising polypeptides and polynucleotides for stimulating
PT the immune system and for treating malignancies associated with
PT overexpression of the HER-2 protein.
XX
PS Claim 1; Page 37; 51pp; English.
XX
CC The present invention relates to compositions for stimulating the immune
CC system and for treating malignancies associated with overexpression of
CC the HER-2 protein. The compositions comprise immunogenic groups of the
CC HER-2 proteins. The present sequence is one such peptide used in the
CC compositions of the present invention. The compositions can be used for
CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers
XX
SQ Sequence 22 AA;
Query Match 100.0%; Score 114; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 INGTSCVDLDDKGCFAEQR 20
DB 1 INGTSCVDLDDKGCFAEQR 20
RESULT 2
ABJ18655
ID ABJ18655 standard; peptide; 30 AA.
XX
AC ABJ18655;
XX
DT 20-FEB-2003 (first entry)
XX
DE ErbB receptor antagonist peptide #38.
XX
KW Vaccine; erbB receptor antagonist; tumour; cancer; gastric cancer;
KW breast cancer; prostate cancer.
```

```
XX Unidentified.
OS
PN WO200281649-A2.
XX
PD 17-OCT-2002.
XX
PF 08-APR-2002; 2002WO-US011211.
XX
PR 06-APR-2001; 2001US-0282037P.
PR 03-AUG-2001; 2001US-0309864P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Greene MI, Zhang H, Murali R, Richter M, Berezov A, Liu Q;
PI Chen J;
XX
DR WPI; 2003-075482/07.
XX
PT New peptide antagonists against erbB receptors, useful for preventing or
PT treating tumors or cancers (e.g. lung adenocarcinomas, breast carcinomas
PT or prostate cancer) in humans.
XX
PS Disclosure; Page 22; 115pp; English.
XX
CC The invention comprises peptide antagonists designed to target erbB
CC receptors. The erbB receptor peptide antagonists are useful for
CC preventing tumours and cancers (e.g. gastric cancer, breast cancer and
CC prostate cancer). The present amino acid sequence represents an erbB
CC receptor peptide antagonist of the invention
XX
SQ Sequence 30 AA;
Query Match 92.1%; Score 105; DB 6; Length 30;
Best Local Similarity 95.0%; Pred. No. 1.8e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 INGTSCVDLDDKGCFAEQR 20
DB 4 INGTSCVDLDDKGCFAEQR 23
RESULT 3
ABP59259
ID ABP59259 standard; peptide; 30 AA.
XX
AC ABP59259;
XX
DT 10-MAY-2003 (first entry)
XX
DE Peptide epitope for generating cell surface receptor antibody, SEQ ID 38.
XX
KW Antibody; cytostatic; vaccine; cell surface receptor; erbB; TNF;
KW immunoglobulin gene superfamily; IgSF; epitope.
XX
OS Synthetic.
XX
PN WO2003012072-A2.
XX
PD 13-FEB-2003.
XX
PF 05-AUG-2002; 2002WO-US024892.
XX
PR 03-AUG-2001; 2001US-0309864P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Greene MI, Zhang H, Richter M, Murali R;
XX
DR WPI; 2003-268117/26.
XX
PT Novel antibody for preventing, treating, imaging or diagnosing tumor, has
PT specificity for activated cell surface receptor and binds specifically to
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:35:58 ; Search time 29.5455 Seconds  
(without alignments)  
55.965 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSCVLDLDDKCPAEQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5.COMB.pep:\*

2: /cgn2\_6/prodata/1/iaa/6.COMB.pep:\*

3: /cgn2\_6/prodata/1/iaa/H.COMB.pep:\*

4: /cgn2\_6/prodata/1/iaa/ECTUS.COMB.pep:\*

5: /cgn2\_6/prodata/1/iaa/RE.COMB.pep:\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	782	1	US-09-146-283-4
2	105	92.1	782	2	US-08-579-823A-4
3	105	92.1	782	2	US-09-344-195-4
4	105	92.1	1255	1	US-08-467-083-68
5	105	92.1	1255	1	US-08-414-417B-68
6	105	92.1	1255	1	US-08-484-438-8
7	105	92.1	1255	1	US-08-486-348A-68
8	105	92.1	1255	1	US-08-625-101-2
9	105	92.1	1255	1	US-08-468-545B-68
10	105	92.1	1255	1	US-08-356-786-2
11	105	92.1	1255	2	US-08-466-680B-68
12	105	92.1	1255	2	US-09-527-487-2
13	105	92.1	1255	2	US-09-811-115-3
14	105	92.1	1255	2	US-09-354-533-68
15	105	92.1	1255	2	US-09-441-411-6
16	105	92.1	1255	2	US-09-167-516-2
17	95	83.3	624	2	US-08-422-108-1
18	95	83.3	624	2	US-08-422-734-1
19	95	83.3	645	2	US-09-602-812A-13
20	50	43.9	122	2	US-09-732-210-241
21	50	43.9	122	2	US-09-732-210-585
22	50	43.9	125	2	US-09-438-185A-639
23	47	41.2	274	2	US-09-188-930-336
24	47	41.2	274	2	US-09-312-283C-336
25	47	41.2	439	2	US-09-409-096-6
26	47	41.2	443	1	US-08-833-963C-2
27	47	41.2	443	2	US-08-980-514-1
28	47	41.2	47	2	US-09-949-016-7792
29	45	39.5	116	2	US-09-270-767-57685
30	45	39.5	124	2	US-09-270-767-31669
31	45	39.5	124	2	US-09-270-767-46886
32	45	39.5	203	2	US-09-248-796A-15633
33	45	39.5	342	2	US-09-902-540-15847
34	44	38.6	334	2	US-09-800-729-109
35	44	38.6	780	2	US-09-248-796A-16702
36	44	38.6	1060	2	US-09-248-796A-14123
37	43	37.7	324	1	US-08-484-938B-22
38	43	37.7	324	1	US-08-484-158B-22
39	43	37.7	324	1	US-08-484-596A-22
40	43	37.7	324	1	US-08-480-150A-22
41	43	37.7	324	2	US-08-458-731-22
42	43	37.7	324	2	US-08-149-223A-22
43	43	37.7	416	2	US-09-538-092-392
44	43	37.7	585	2	US-09-270-767-59685
45	43	37.7	621	2	US-09-489-039A-10378
46	43	37.7	801	2	US-09-270-767-44264
47	42.5	37.3	201	2	US-09-710-279-1514
48	42.5	37.3	255	2	US-09-134-001C-3498
49	42	36.8	90	2	US-09-248-796A-19142
50	42	36.8	101	2	US-09-716-129-80
51	42	36.8	132	2	US-09-716-129-179
52	42	36.8	321	2	US-09-270-767-33762
53	42	36.8	321	2	US-09-270-767-48979
54	42	36.8	324	2	US-09-605-703B-1888
55	42	36.8	1390	2	US-09-902-540-11251
56	42	36.8	2813	2	US-08-896-449A-2
57	42	36.8	2813	2	US-09-132-652-2
58	42	36.8	2813	2	US-09-886-900A-2
59	42	36.8	2813	2	US-09-662-478C-2
60	42	36.8	3457	1	US-08-416-603-4
61	41	36.0	136	2	US-09-252-991A-16772
62	41	36.0	864	1	US-08-209-521-29
63	41	36.0	864	2	US-09-265-503B-138
64	40.5	35.5	497	2	US-09-252-991A-31534
65	40.5	35.5	726	6	5208144-37
66	40	35.1	120	2	US-09-248-796A-19679
67	40	35.1	124	2	US-09-513-999C-4218
68	40	35.1	214	2	US-09-902-540-16784
69	40	35.1	232	2	US-09-248-796A-14122
70	40	35.1	234	2	US-10-000-489-20
71	40	35.1	246	2	US-09-248-796A-18240
72	40	35.1	265	2	US-09-252-991A-24304
73	40	35.1	303	6	5340934-13
74	40	35.1	317	2	US-09-270-767-43541
75	40	35.1	368	2	US-09-668-097A-38
76	40	35.1	551	2	US-09-252-991A-26416
77	39.5	34.6	194	2	US-08-679-493A-215
78	39.5	34.6	194	2	US-08-679-493A-216
79	39.5	34.6	357	2	US-09-645-168-3
80	39.5	34.6	1172	1	US-08-313-288B-19
81	39.5	34.6	1172	2	US-09-949-016-6333
82	39.5	34.6	2594	2	US-08-718-388-7
83	39.5	34.6	4654	2	US-08-476-515A-84
84	39.5	34.6	4655	2	US-08-652-877-84
85	39.5	34.6	4655	2	US-08-652-877-86
86	39.5	34.6	4655	2	US-08-652-877-88
87	39.5	34.6	4655	2	US-08-652-877-90
88	39.5	34.6	5405	2	US-08-718-388-9
89	39	34.2	187	2	US-08-463-682-1
90	39	34.2	213	6	5212074-6
91	39	34.2	220	2	US-08-463-682-24
92	39	34.2	226	1	US-08-463-682-23
93	39	34.2	226	1	US-08-347-594A-4
94	39	34.2	227	2	US-09-252-991A-24444
95	39	34.2	229	2	US-09-248-796A-16698
96	39	34.2	230	2	US-09-902-540-15990
97	39	34.2	240	2	US-09-461-912A-41
98	39	34.2	247	1	US-08-797-689-4
99	39	34.2	247	2	US-09-984-186-4
100	39	34.2	257	2	US-09-949-016-11037

## ALIGNMENTS

```
RESULT 1
US-09-146-283-4
; Sequence 4, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4
Query Match 92.1%; Score 105; DB 1; Length 782;
Best Local Similarity 95.0%; Pred. No. 1.2e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 INGTHSCVDLDDKGCPEQR 20
Db 628 INCTHSCVDLDDKGCPEQR 647

RESULT 2
US-09-579-823A-4
; Sequence 4, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
```

```
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4
Query Match 92.1%; Score 105; DB 2; Length 782;
Best Local Similarity 95.0%; Pred. No. 1.2e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 INGTHSCVDLDDKGCPEQR 20
Db 628 INCTHSCVDLDDKGCPEQR 647

RESULT 3
US-09-344-195-4
; Sequence 4, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:50:00 ; Search time 99.5454 Seconds  
(without alignments)  
83.947 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSCVDLDDKGCPAEQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 segs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications\_AA\_Main:\*

1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:\*

3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:\*

4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep:\*

5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep:\*

6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	493	4	US-10-412-804A-2
2	105	92.1	653	3	US-09-854-356-3
3	105	92.1	678	5	US-10-956-373-2
4	105	92.1	680	5	US-10-846-113A-5
5	105	92.1	690	4	US-10-412-804A-11
6	105	92.1	712	3	US-09-854-356-7
7	105	92.1	715	4	US-10-412-804A-10
8	105	92.1	717	5	US-10-956-373-14
9	105	92.1	720	4	US-10-412-804A-6
10	105	92.1	919	3	US-09-854-356-6
11	105	92.1	1253	4	US-10-146-473-72
12	105	92.1	1255	3	US-09-811-123-9
13	105	92.1	1255	3	US-09-811-115-3
14	105	92.1	1255	3	US-09-765-973-2
15	105	92.1	1255	3	US-09-354-533-68
16	105	92.1	1255	3	US-09-769-508-2
17	105	92.1	1255	3	US-09-854-356-1
18	105	92.1	1255	3	US-09-930-125-2
19	105	92.1	1255	3	US-09-441-411-6
20	105	92.1	1255	3	US-09-984-092-4
21	105	92.1	1255	4	US-10-207-655-45
22	105	92.1	1255	4	US-10-177-293-126
23	105	92.1	1255	4	US-10-207-498-6
24	105	92.1	1255	4	US-10-338-730-2
25	105	92.1	1255	4	US-10-313-644-2
26	105	92.1	1255	4	US-10-322-892-4
27	105	92.1	1255	4	US-10-272-437A-28

28	105	92.1	1355	4	US-10-117-937-594	Sequence 594, App
29	105	92.1	1355	4	US-10-418-027-3	Sequence 3, Appli
30	105	92.1	1355	4	US-10-394-322A-17	Sequence 17, Appl
31	105	92.1	1355	4	US-10-245-871-553	Sequence 553, App
32	105	92.1	1355	4	US-10-435-696-36	Sequence 36, Appl
33	105	92.1	1355	4	US-10-149-138-4641	Sequence 4641, Ap
34	105	92.1	1355	4	US-10-469-162-3	Sequence 3, Appli
35	105	92.1	1355	4	US-10-253-286-553	Sequence 553, App
36	105	92.1	1355	4	US-10-647-005-68	Sequence 68, Appl
37	105	92.1	1355	4	US-10-149-138-4641	Sequence 4641, Ap
38	105	92.1	1355	4	US-10-441-779C-4	Sequence 4, Appli
39	105	92.1	1355	4	US-10-734-564-126	Sequence 126, App
40	105	92.1	1355	4	US-10-657-022-90	Sequence 90, Appl
41	105	92.1	1355	5	US-10-762-128-6	Sequence 6, Appli
42	105	92.1	1355	5	US-10-484-067-1	Sequence 1, Appli
43	105	92.1	1355	5	US-10-723-860-9	Sequence 9, Appli
44	105	92.1	1355	5	US-10-615-343-17	Sequence 17, Appl
45	105	92.1	1355	5	US-10-794-514A-1	Sequence 1, Appli
46	105	92.1	1355	5	US-10-871-708-9	Sequence 9, Appli
47	105	92.1	1355	5	US-10-983-340-17	Sequence 17, Appl
48	105	92.1	1355	5	US-10-846-113A-27	Sequence 27, Appl
49	105	92.1	1355	6	US-11-067-064-594	Sequence 594, App
50	105	92.1	1355	6	US-11-121-347-68	Sequence 68, Appl
51	105	92.1	1355	6	US-11-037-713-12	Sequence 12, Appl
52	105	92.1	1355	6	US-11-067-159-594	Sequence 594, App
53	105	92.1	1294	5	US-10-956-373-10	Sequence 10, Appl
54	98	86.0	177	5	US-10-949-667-5	Sequence 5, Appli
55	98	86.0	654	3	US-09-854-356-8	Sequence 8, Appli
56	98	86.0	1256	3	US-09-854-356-2	Sequence 2, Appli
57	98	86.0	1256	3	US-09-854-356-14	Sequence 14, Appl
58	98	86.0	1257	5	US-10-484-067-2	Sequence 2, Appli
59	98	86.0	1260	3	US-09-870-759-118	Sequence 118, App
60	98	86.0	1260	3	US-09-751-708A-118	Sequence 118, App
61	98	86.0	1260	4	US-10-428-817A-114	Sequence 114, App
62	98	86.0	1260	5	US-10-937-758A-95	Sequence 95, Appl
63	98	86.0	1260	5	US-10-949-667-8	Sequence 8, Appli
64	95	83.3	645	3	US-09-921-161-1	Sequence 1, Appli
65	95	83.3	645	4	US-10-268-501-13	Sequence 13, Appl
66	95	83.3	645	4	US-10-608-626-13	Sequence 13, Appl
67	95	83.3	645	5	US-10-719-310-13	Sequence 13, Appl
68	95	83.3	645	5	US-11-044-749-13	Sequence 13, Appl
69	95	83.3	645	6	US-11-154-465-13	Sequence 13, Appl
70	88	77.2	15	4	US-10-149-138-3948	Sequence 3948, Ap
71	88	77.2	15	4	US-10-149-138-4529	Sequence 4529, Ap
72	88	77.2	15	4	US-10-149-138-4529	Sequence 4529, Ap
73	88	77.2	15	4	US-10-149-137A-136	Sequence 136, App
74	88	77.2	15	4	US-10-484-067-10	Sequence 10, Appl
75	82	71.9	275	5	US-10-149-138-3734	Sequence 3734, Ap
76	81	71.1	15	4	US-10-149-138-3734	Sequence 3734, Ap
77	81	71.1	15	4	US-10-149-138-3734	Sequence 11, Appl
78	81	71.1	16	3	US-09-811-123-11	Sequence 11, Appl
79	81	71.1	50	5	US-10-484-067-11	Sequence 11, Appl
80	78	68.4	50	5	US-10-484-067-5	Sequence 5, Appli
81	78	68.4	275	5	US-10-484-067-4	Sequence 4, Appli
82	66	57.9	11	4	US-10-149-138-194	Sequence 194, App
83	66	57.9	11	4	US-10-149-138-2511	Sequence 2511, Ap
84	66	57.9	11	4	US-10-149-138-194	Sequence 194, App
85	66	57.9	11	4	US-10-149-138-2511	Sequence 2511, Ap
86	62	54.4	11	4	US-10-149-138-2798	Sequence 2798, Ap
87	62	54.4	11	4	US-10-149-138-3404	Sequence 3404, Ap
88	62	54.4	11	4	US-10-149-138-3404	Sequence 3404, Ap
89	62	54.4	11	4	US-10-149-138-3841	Sequence 3841, Ap
90	59	51.8	15	4	US-10-149-138-3841	Sequence 3841, Ap
91	59	51.8	15	4	US-10-149-138-3841	Sequence 3841, Ap
92	58	50.9	10	4	US-10-149-138-2521	Sequence 2521, Ap
93	58	50.9	10	4	US-10-149-138-3233	Sequence 3233, Ap
94	58	50.9	10	4	US-10-149-138-2521	Sequence 2521, Ap
95	58	50.9	10	4	US-10-149-138-3233	Sequence 3233, Ap
96	58	50.9	11	4	US-10-149-138-2522	Sequence 2522, Ap
97	58	50.9	11	4	US-10-149-138-2522	Sequence 2522, Ap
98	57	50.0	10	4	US-10-149-138-3106	Sequence 3106, Ap
99	57	50.0	10	4	US-10-149-138-3106	Sequence 3106, Ap
100	57	50.0	11	4	US-10-149-138-859	Sequence 859, App

Db 630 INCTHSCVDLDDKGCPAEQR 649

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 3, 2005, 07:48:55 ; Search time 5.90909 Seconds  
(without alignments)  
16.207 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTHTSCVDLDKGFPAQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications\_AA\_New.\*

1: /cgn2\_6/ptodata1/pubpaa/US09\_NEW\_PUB.pep.\*

2: /cgn2\_6/ptodata1/pubpaa/US06\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata1/pubpaa/US07\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata1/pubpaa/US08\_NEW\_PUB.pep.\*

5: /cgn2\_6/ptodata1/pubpaa/PT\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata1/pubpaa/US10\_NEW\_PUB.pep.\*

7: /cgn2\_6/ptodata1/pubpaa/US11\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	1255	7	US-11-022-562-213
2	47	41.2	443	6	US-10-131-826A-318
3	42.5	37.3	201	6	US-10-793-626-1514
4	39.5	34.6	251	6	US-10-485-517-167
5	39.5	34.6	5405	7	US-11-108-172-1116
6	39	34.2	2919	6	US-10-821-234-1133
7	38.5	33.8	703	6	US-10-821-234-1133
8	38	33.3	143	6	US-10-821-234-1216
9	38	33.3	353	6	US-10-131-826A-296
10	38	33.3	354	6	US-10-467-657-7260
11	38	33.3	390	6	US-10-858-730-59
12	38	33.3	1042	7	US-11-067-811-1
13	38	33.3	1113	7	US-11-067-811-4
14	37.5	32.9	420	6	US-10-131-826A-290
15	37.5	32.9	1437	7	US-11-074-176-96
16	37	32.5	104	7	US-11-065-669-5
17	37	32.5	402	7	US-11-051-568-29
18	37	32.5	548	7	US-11-137-465-47
19	37	32.5	853	6	US-10-420-192-6
20	36.5	32.0	509	6	US-10-821-234-1093
21	36.5	32.0	1614	7	US-11-108-528-82
22	36.5	32.0	3002	6	US-10-821-234-916
23	36	31.6	304	6	US-10-467-657-7616
24	36	31.6	898	6	US-10-624-932-2
25	36	31.6	1548	7	US-11-108-172-1095
26	35.5	31.1	989	6	US-10-821-234-975
27	35.5	31.1	997	7	US-11-113-424-37
28	35	30.7	146	6	US-10-467-657-3166
29	35	30.7	246	6	US-10-485-517-361
30	35	30.7	275	7	US-11-113-424-79
31	35	30.7	346	6	US-10-131-826A-432
32	35	30.7	346	6	US-10-967-457-77
33	35	30.7	621	6	US-10-821-234-1376
34	35	30.7	622	7	US-11-021-441-35
35	35	30.7	852	6	US-10-645-441-15
36	35	30.7	2847	6	US-10-821-234-1303
37	34.5	30.3	161	7	US-11-113-424-38
38	34.5	30.3	387	7	US-11-137-465-57
39	34.5	30.3	427	6	US-10-508-263-96
40	34.5	30.3	897	7	US-11-137-465-35
41	34.5	30.3	961	7	US-11-113-424-35
42	34.5	30.3	964	7	US-11-137-465-58
43	34.5	30.3	965	7	US-11-113-424-2
44	34.5	30.3	965	7	US-11-147-047-51
45	34.5	30.3	993	7	US-11-137-465-36
46	34.5	30.3	999	7	US-11-113-424-36
47	34	29.8	112	6	US-10-821-234-1236
48	34	29.8	153	6	US-10-793-626-3170
49	34	29.8	196	6	US-10-967-527A-27
50	34	29.8	370	6	US-10-821-234-1502
51	34	29.8	404	6	US-10-467-657-9211
52	34	29.8	1442	6	US-10-793-626-2052
53	34	29.8	2515	7	US-11-113-424-53
54	33.5	29.4	172	7	US-11-177-010-2
55	33.5	29.4	172	7	US-11-177-010-4
56	33.5	29.4	197	7	US-11-055-822-908
57	33.5	29.4	419	6	US-10-821-234-1556
58	33.5	29.4	1615	7	US-11-108-528-80
59	33.5	29.4	5179	7	US-11-108-172-1068
60	33	28.9	124	6	US-10-467-657-7022
61	33	28.9	131	6	US-10-467-657-3792
62	33	28.9	140	6	US-10-467-657-1360
63	33	28.9	160	6	US-10-467-657-2884
64	33	28.9	176	7	US-11-055-822-210
65	33	28.9	176	7	US-11-055-822-726
66	33	28.9	292	6	US-10-467-657-3894
67	33	28.9	292	6	US-10-467-657-7242
68	33	28.9	345	6	US-10-793-626-3168
69	33	28.9	346	6	US-10-793-626-2034
70	33	28.9	361	7	US-11-012-762-42
71	33	28.9	409	7	US-11-102-240-22
72	33	28.9	496	6	US-10-793-626-1386
73	33	28.9	557	7	US-11-191-374-4
74	33	28.9	557	7	US-11-191-375-4
75	33	28.9	639	7	US-11-113-837-21
76	33	28.9	673	7	US-11-102-240-16
77	33	28.9	683	6	US-10-821-234-1314
78	33	28.9	683	7	US-11-081-566-2
79	33	28.9	683	7	US-11-108-172-1122
80	33	28.9	697	7	US-11-082-389-362
81	33	28.9	711	6	US-10-510-386-10
82	33	28.9	741	7	US-11-191-374-18
83	33	28.9	741	7	US-11-191-375-18
84	33	28.9	742	7	US-11-191-374-46
85	33	28.9	742	7	US-11-191-375-46
86	33	28.9	941	6	US-10-131-826A-464
87	33	28.9	950	6	US-10-467-657-854
88	33	28.9	956	7	US-11-113-424-39
89	33	28.9	1044	7	US-11-091-668-2
90	33	28.9	1210	7	US-11-191-374-10
91	33	28.9	1210	7	US-11-191-375-10
92	33	28.9	1221	6	US-10-858-730-232
93	33	28.9	1313	7	US-11-091-668-4
94	33	28.9	1400	6	US-10-821-234-1045
95	32.5	28.5	105	6	US-10-793-626-158
96	32.5	28.5	193	7	US-11-085-775-4
97	32.5	28.5	203	6	US-10-467-657-3082
98	32.5	28.5	747	7	US-11-018-018-1

Sequence 975, Appl  
Sequence 37, Appl  
Sequence 3166, Ap  
Sequence 361, App  
Sequence 79, Appl  
Sequence 432, App  
Sequence 77, Appl  
Sequence 1376, Ap  
Sequence 35, Appl  
Sequence 15, Appl  
Sequence 1303, Ap  
Sequence 38, Appl  
Sequence 57, Appl  
Sequence 96, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 2, Appl  
Sequence 51, Appl  
Sequence 36, Appl  
Sequence 1236, Ap  
Sequence 3170, Ap  
Sequence 27, Appl  
Sequence 1502, Ap  
Sequence 9211, Ap  
Sequence 2052, Ap  
Sequence 53, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 908, App  
Sequence 1556, Ap  
Sequence 80, Appl  
Sequence 1068, Ap  
Sequence 7022, Ap  
Sequence 1260, Ap  
Sequence 3792, Ap  
Sequence 2884, Ap  
Sequence 210, App  
Sequence 726, App  
Sequence 3894, Ap  
Sequence 7242, Ap  
Sequence 3168, Ap  
Sequence 2034, Ap  
Sequence 42, Appl  
Sequence 22, Appl  
Sequence 1386, Ap  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 21, Appl  
Sequence 16, Appl  
Sequence 1314, Ap  
Sequence 2, Appl  
Sequence 122, App  
Sequence 362, App  
Sequence 10, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 46, Appl  
Sequence 464, App  
Sequence 854, App  
Sequence 39, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 222, App  
Sequence 4, Appl  
Sequence 1045, Ap  
Sequence 158, App  
Sequence 4, Appl  
Sequence 3082, Ap  
Sequence 1, Appl

99 32.5 28.5 747 7 US-11-047-757-1 Sequence 1, Appli  
100 32.5 28.5 1613 7 US-11-108-528-84 Sequence 84, Appli

## ALIGNMENTS

RESULT 1  
US-11-022-562-213  
; Sequence 213, Application US/11022562  
; Publication No. US20050249742A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruprecht, Ruth M.  
; APPLICANT: Shisong, Jiang  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING  
; TITLE OF INVENTION: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE  
; FILE REFERENCE: DFN-043CN  
; CURRENT APPLICATION NUMBER: US/11/022,562  
; CURRENT FILING DATE: 2004-12-22  
; PRIOR APPLICATION NUMBER: PCT/US03/20322  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 60/392718  
; PRIOR FILING DATE: 2002-06-27  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 213  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-022-562-213

Query Match 92.1%; Score 105; DB 7; Length 1255;  
Best Local Similarity 95.0%; Pred. No. 1.1e-09;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 INGTSCVDLDDKGPAPQR 20  
DB 628 INCTHSCVDLDDKGPAPQR 647

RESULT 2  
US-10-131-826A-318  
; Sequence 318, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C128  
; CURRENT APPLICATION NUMBER: US/10/131,826A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 318  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-131-826A-318

Query Match 41.2%; Score 47; DB 6; Length 443;  
Best Local Similarity 58.8%; Pred. No. 0.85;  
Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;  
QY 6 SCVDLD--DKGCPAPQR 20  
DB 200 SCVDVNECDMGAPCEQR 216

RESULT 3  
US-10-793-626-1514  
; Sequence 1514, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1514  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-10-793-626-1514

Query Match 37.3%; Score 42.5; DB 6; Length 201;  
Best Local Similarity 43.5%; Pred. No. 2.1;  
Matches 10; Conservative 2; Mismatches 6; Indels 5; Gaps 2;  
QY 2 NGTHSCVDLDDKGC---PAEQR 20  
DB 48 NGVHTCID-TSAGCVNDTPAFNR 69

RESULT 4  
US-10-485-517-167  
; Sequence 167, Application US/10485517  
; Publication No. US20050256299A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Biosynexus Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:36:13 ; Search time 20.9091 Seconds  
(without alignments)  
92.033 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSHCVLDLDKGCPEQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR\_80:\*

1: pirl:\*

2: pirl:\*

3: pirl:\*

4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	1255	1 A24571	protein-tyrosine k
2	98	86.0	1254	2 I48161	p-185 precursor -
3	98	86.0	1250	1 TVRNU	protein-tyrosine k
4	51	44.7	235	2 S30001	ribosomal protein
5	50	43.9	122	2 B86570	L14 ribosomal prot
6	50	43.9	122	2 G72054	ribosomal protein
7	50	43.9	122	2 E81664	ribosomal protein
8	50	43.9	122	2 D42645	ribosomal protein
9	49	43.0	303	2 A84344	NADH dehydrogenase
10	47	41.2	1435	2 S54697	DNA polymerase III
11	47	41.2	1435	2 C90596	hypothetical prote
12	46.5	40.8	548	2 T16642	hypothetical prote
13	46	40.4	462	1 A37986	interleukin-6 rece
14	45	39.5	111	2 T36163	probable transcript
15	45	39.5	140	2 C90850	conserved hypotet
16	44	38.6	124	2 G87028	hypothetical prote
17	44	38.6	129	2 AE0206	conserved hypotet
18	44	38.6	542	2 B90090	very similar to mo
19	44	38.6	747	2 T23042	hypothetical prote
20	44	38.6	749	2 T23045	hypothetical prote
21	44	38.6	1419	2 T30531	agglutinin-like ad
22	44	38.6	2643	2 T29149	hypothetical prote
23	43.5	38.2	4660	2 T42737	gp330 protein prec
24	43	37.7	304	2 G85068	N7-like protein [i
25	43	37.7	416	2 S48957	hypothetical prote
26	43	37.7	490	2 AC2458	hypothetical prote
27	43	37.7	2907	2 A57278	fibrillin-2 precur
28	43	37.7	2918	2 A54105	fibrillin-2 precur
29	42	36.8	99	2 H72868	AcOrf-150 protein

30	42	36.8	174	2 H96001	hypothetical prote
31	42	36.8	232	2 S56275	probable membrane
32	42	36.8	440	2 T24323	hypothetical prote
33	42	36.8	516	2 AH2417	hypothetical prote
34	42	36.8	1462	1 B36182	protein-tyrosine-p
35	42	36.8	4543	1 A53102	alpha-2-macroglobu
36	41.5	36.4	1847	2 T18308	probable vitelloge
37	41.5	36.4	1984	2 T13171	probable vitelloge
38	41	36.0	108	2 B87095	ferredoxin (impor
39	41	36.0	108	2 A70876	probable ferredoxi
40	41	36.0	115	2 T41887	AcMNPV orf150 - Bo
41	41	36.0	283	2 S75226	esterase elr1916 -
42	41	36.0	291	2 H97090	thioredoxin reduct
43	41	36.0	366	2 D45558	epidermal growth f
44	41	36.0	387	2 JE0360	gamma-Butyrobetain
45	41	36.0	414	2 PS0323	von Willebrand fac
46	41	36.0	424	2 B38176	samb protein - Sal
47	41	36.0	506	2 S71591	aspartic proteinas
48	41	36.0	1401	2 S77657	cyclic peptide syn
49	41	36.0	1717	1 A45558	epidermal growth f
50	41	36.0	2150	2 T32497	hypothetical prote
51	40.5	35.5	341	2 T19224	hypothetical prote
52	40.5	35.5	770	2 T00203	LDL receptor-relat
53	40.5	35.5	1297	2 T30274	protosolaisin - se
54	40.5	35.5	3871	2 T22812	hypothetical prote
55	40	35.1	105	2 S08287	ferredoxin [4Fe-4S
56	40	35.1	106	1 FEMVFS	ferredoxin [3Fe-4S
57	40	35.1	121	2 G64152	hypothetical prote
58	40	35.1	136	2 A85711	probable IS encode
59	40	35.1	195	2 AC0123	probable exported
60	40	35.1	206	2 D91109	hypothetical prote
61	40	35.1	261	2 D83101	conserved hypotet
62	40	35.1	261	2 G85954	unknown protein en
63	40	35.1	297	2 S04455	hypothetical prote
64	40	35.1	303	1 GEHUN	osteonectin precur
65	40	35.1	304	1 GEBON	osteonectin precur
66	40	35.1	304	2 S04456	hypothetical prote
67	40	35.1	330	2 S60211	fomB protein - Str
68	40	35.1	348	2 T29288	hypothetical prote
69	40	35.1	354	2 T22274	hypothetical prote
70	40	35.1	392	2 H97026	flavoprotein limpo
71	40	35.1	427	2 G00039	von Willebrand fac
72	40	35.1	434	2 T21995	hypothetical prote
73	40	35.1	463	2 B91112	hypothetical prote
74	40	35.1	468	2 H72230	glutamate synthase
75	40	35.1	468	2 T50873	hypothetical prote
76	40	35.1	512	2 B90670	hypothetical prote
77	40	35.1	512	2 H90766	hypothetical prote
78	40	35.1	512	2 C90792	hypothetical prote
79	40	35.1	512	2 D90836	hypothetical prote
80	40	35.1	512	2 B90856	hypothetical prote
81	40	35.1	512	2 H90906	hypothetical prote
82	40	35.1	512	2 G90977	hypothetical prote
83	40	35.1	512	2 C91197	hypothetical prote
84	40	35.1	512	2 A85653	unknown protein en
85	40	35.1	512	2 A86044	hypothetical prote
86	40	35.1	614	2 T19963	hypothetical prote
87	40	35.1	667	2 S48285	probable glycine-t
88	40	35.1	1052	2 T50127	hypothetical prote
89	40	35.1	1064	2 H64516	hypothetical prote
90	40	35.1	1260	2 S60896	agglutinin-like pr
91	40	35.1	1437	2 F96783	unknown protein P2
92	40	35.1	2233	2 T28669	surface protein 51
93	39.5	34.6	129	2 A72606	hypothetical prote
94	39.5	34.6	251	2 F98785	formate acetyltran
95	39.5	34.6	357	2 S38338	naringenin 3-dioxy
96	39.5	34.6	358	2 T45755	flavanone 3-hydrox
97	39.5	34.6	1172	1 TSHUP2	thrombospondin 2 p
98	39.5	34.6	1172	2 A42587	thrombospondin 2 p
99	39.5	34.6	2180	2 T29764	hypothetical prote
100	39	34.2	114	2 E70759	probable foxA prot

## ALIGNMENTS

```
RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 05-Oct-2004
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: UNIPROT:P04626; UNIPARC:UPI000003F55F; GB:X03363; NID:g31197; PIDN:C
R:Semba, K.; Kanata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: UNIPARC:UPI000016A8A7; GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:
R:Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromoe
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: UNIPARC:UPI000016AA26; GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALH', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: UNIPARC:UPI00001725C7; GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: UNIPARC:UPI0000070A3F; GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>
A:Cross-references: UNIPARC:UPI0000000427; GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:
C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C:Genetics:
A:Gene: GDB:ERBB2; NGL; NEU; HER-2
A:Cross-references: GDB:L20613; OMIM:164870
A:Map position: 17q21.1-17q21.1
A:Introns: 25/1; 75/3; 147/1; 883/3
A>Note: the list of introns is incomplete
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F:22-653/Domain: extracellular #status predicted <EXT>
F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
```

```
F:654-675/Domain: transmembrane #status predicted <TMM>
F:676-1255/Domain: intracellular #status predicted <INT>
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif
F:68,124,187,259,530,571,629/Binding site: carboxylate (Asn) (covalent) #status predict
F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:753/Active site: Lys #status predicted
F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 92.1%; Score 105; DB 1; Length 1255;
Best Local Similarity 95.0%; Pred. No. 5.7e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 INGTHSCVDLDDKGCPEAQR 20
|||
Db 628 INCTHSCVDLDDKGCPEAQR 647

RESULT 2
I48161
p-185 precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004
C:Accession: I48161
R:Nakamura, T.; Uehijima, T.; Ishizaka, T.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa,
Gene 140, 251-255, 1994
A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
A:Reference number: I48161; MUID:94193007; PMID:7908275
A:Accession: I48161
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1254 <RES>
A:Cross-references: UNIPROT:Q60553; UNIPARC:UPI000012A111; GB:D16295; NID:g4932336; PIDN:
C:Genetics:
A:Gene: neu
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:Keywords: ATP
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif

Query Match 86.0%; Score 98; DB 2; Length 1254;
Best Local Similarity 85.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 INGTHSCVDLDDKGCPEAQR 20
|||
Db 628 INCTHSCVDLDDKGCPEAQR 647

RESULT 3
TVRTNU
Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004
C:Accession: A24562; A61204
R:Barbmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A:Reference number: A24562; MUID:86118662; PMID:3945311
A:Accession: A24562
A:Molecule type: mRNA
A:Residues: 1-1260 <BAR>
A:Cross-references: UNIPROT:P06494; UNIPARC:UPI0000161883; EMBL:X03362; NID:g56745; PIDN:
R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
Carcinogenesis 12, 1975-1978, 1991
A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no r
2-thiazolylformamide or N-methyl-N-nitrosourea.
A:Reference number: A61204; MUID:92035293; PMID:1682063
A:Accession: A61204
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 637-663, 'V', 665-702 <MAS>
A:Cross-references: UNIPARC:UPI00001725C8
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 07:35:08 ; Search time 132.727 Seconds  
(without alignments)  
106.312 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSCVDLDDRGCPAEQR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Uniprot\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	435	2 Q62MM4_HUMAN	P04626 homo sapien
2	105	92.1	1255	1 ERBB2_HUMAN	P04626 homo sapien
3	98	86.0	176	2 Q92V55_RAT	Q923v5 rattus norv
4	98	86.0	711	2 Q80Y89_MOUSE	Q80Y89 mus musculu
5	98	86.0	881	2 Q8C0E7_MOUSE	Q8C0E7 m mus muscu
6	98	86.0	1254	1 ERBB2_MESAU	P06494 rattus norv
7	98	86.0	1257	1 ERBB2_RAT	P06494 rattus norv
8	98	86.0	1259	2 Q8K3F9_RAT	Q8K3f9 rattus norv
9	98	86.0	1305	2 Q6ZPE0_MOUSE	Q6ZPE0 mus musculu
10	97	85.1	1259	1 ERBB2_CANFA	Q18735 canis fami
11	61	53.5	311	2 Q41LT0_GIBZE	Q41lt0 gibberella
12	54.5	47.8	1275	2 Q5EBY4_BRARE	Q5eb4 brachydanio
13	53	46.5	293	2 Q7YQX5_PHATA	Q7Yqx5 phascogale
14	53	46.5	305	2 Q7YQX4_9META	Q7Yqx4 sminthopsis
15	53	46.5	314	2 Q7YQX3_DASAL	Q7Yqx3 daeyurus al
16	53	46.5	322	2 Q7YQY0_DRONU	Q7Yqy0 dromiclops
17	51	44.7	235	1 RS6_KLUMA	P41798 kluyveromyc
18	51	44.7	287	2 Q9FE11_9POAL	Q9fe11 avena pilos
19	51	44.7	287	2 Q9FEJ2_9POAL	Q9fe12 avena claud
20	51	44.7	287	2 Q9FEJ3_9POAL	Q9fe13 avena ventr
21	51	44.7	1335	2 Q4RG29_TETNG	Q4rg29 tetraodon n
22	50	43.9	122	1 RL14_CHLMU	Q9Pjm4 chlamydia m
23	50	43.9	122	1 RL14_CHLPN	Q927r7 chlamydia p
24	50	43.9	122	1 RL14_CHLTR	P28533 chlamydia t
25	50	43.9	236	1 RS6_KLULA	Q6cm04 kluyveromyc
26	50	43.9	972	2 Q754G9_ASHGO	Q754g9 ashbya gos
27	49.5	43.4	805	2 Q9PTY3_PAROL	Q9pty3 paralicthy
28	49	43.0	243	2 Q86DY7_SCHJA	Q86dy7 schistosoma
29	49	43.0	303	2 Q9HNW3_HALSA	Q9hnw3 halobacteri
30	49	43.0	358	2 Q29134_TUPGL	Q29134 tupaia glis
31	49	43.0	358	2 Q9TV39_TUPGL	Q9tv39 tupaia glis

32	49	43.0	409	2	Q28549_ORYAF	Q28549 orycteropu
33	49	43.0	766	2	Q4P734_USTMA	Q4p734 ustilago ma
34	48	42.1	140	2	Q4TCV6_TETNG	Q4tcv6 tetraodon n
35	48	42.1	145	2	Q9F560_ECOLI	Q9f560 escherichia
36	48	42.1	287	2	Q9FE15_9POAL	Q9fe15 avena macro
37	48	42.1	323	2	Q7YQX9_NORTY	Q7yqx9 notoryctes
38	48	42.1	372	2	Q28449_LOXAF	Q28449 loxodonta a
39	48	42.1	388	2	Q4RIL8_TETNG	Q4ril8 tetraodon a
40	48	42.1	411	2	Q28404_ELEMA	Q28404 elephas max
41	48	42.1	411	2	Q8WMS8_TRIMA	Q8wms8 trichechus
42	48	42.1	425	2	Q28759_PROCA	Q28759 procavia ca
43	48	42.1	427	2	Q28352_DUGDU	Q28352 dugong dugo
44	47	41.2	220	2	Q59ZS3_CANAL	Q59z53 candida alb
45	47	41.2	299	2	Q6QXH9_GVAS	Q6qxh9 agrotis seg
46	47	41.2	319	2	Q8HXM6_9EUTH	Q8hxm6 uropeilus s
47	47	41.2	344	2	Q7YQY1_PHACI	Q7yqy1 phascolarct
48	47	41.2	344	2	Q866Z1_VOMOR	Q866z1 vomphatus ur
49	47	41.2	353	2	Q9SP47_EPIPO	Q9sp47 epiphys po
50	47	41.2	402	2	Q5X184_RAT	Q5x184 rattus norv
51	47	41.2	443	1	FBLN4_CRIGR	O55058 cricetus
52	47	41.2	443	1	FBLN4_HUMAN	O95957 homo sapien
53	47	41.2	443	1	FBLN4_MOUSE	O9wvj9 mus musculu
54	47	41.2	443	2	Q6FH22_HUMAN	Q6fh22 homo sapien
55	47	41.2	443	2	Q96TF5_HUMAN	Q96tf5 homo sapien
56	47	41.2	443	2	Q9H3D5_HUMAN	Q9h3d5 homo sapien
57	47	41.2	443	2	Q542X5_MOUSE	O542x5 mus musculu
58	47	41.2	443	2	Q9JM06_MOUSE	O9jmo6 mus musculu
59	47	41.2	534	2	Q57TS0_9TRYP	O57ts0 trypanosoma
60	47	41.2	542	2	Q8S7G1_ORYSA	Q8s7g1 oryza sativ
61	47	41.2	620	2	Q59G18_HUMAN	O59g18 homo sapien
62	47	41.2	1435	1	DP03_MYCPU	P47729 mycoplasma
63	47	41.2	2691	1	Q4QFA9_LEIMA	Q4qfa9 leishmania
64	46.5	40.8	146	2	Q42178_ARATH	Q42178 arabidopsis
65	46.5	40.8	161	2	Q96FH9_HUMAN	Q96fh9 homo sapien
66	46.5	40.8	260	2	Q5VXB4_HUMAN	Q5vxb4 homo sapien
67	46.5	40.8	349	2	Q9NVG4_HUMAN	Q9nv94 homo sapien
68	46.5	40.8	365	2	Q4URH9_XANCP	Q4urh9 xanthomonas
69	46.5	40.8	365	2	Q8PC19_XANCP	O8pc19 xanthomonas
70	46.5	40.8	370	2	Q6FWK6_CANGA	O6fwk6 candida gla
71	46.5	40.8	494	2	Q8IWV5_HUMAN	O8iww5 homo sapien
72	46.5	40.8	548	2	Q21629_CAEEL	Q21629 caenorhabdi
73	46.5	40.8	549	2	Q9HOD0_HUMAN	Q9hod0 homo sapien
74	46.5	40.8	572	2	Q9HC49_HUMAN	Q9hc49 homo sapien
75	46.5	40.8	682	2	Q8IWV6_HUMAN	O8iww6 homo sapien
76	46.5	40.8	682	2	Q9H988_HUMAN	O9h988 homo sapien
77	46.5	40.8	683	2	Q9H012_HUMAN	O9h012 homo sapien
78	46.5	40.8	683	2	Q8WVK9_HUMAN	O8wvk9 homo sapien
79	46	40.4	122	2	Q5L710_CHLAD	Q5l710 chlamydomphi
80	46	40.4	122	2	Q824P1_CHLCV	Q824p1 chlamydomphi
81	46	40.4	236	1	RS6_DBBHA	O6bkh8 debaryomyce
82	46	40.4	284	1	Q61G31_CAEER	O61g31 caenorhabdi
83	46	40.4	327	1	EXOB_RHILT	O59745 rhizobium l
84	46	40.4	385	2	Q91YC1_9RODE	Q91yc1 phodopus ro
85	46	40.4	385	2	Q91YC6_9RODE	O91yc6 myospalax s
86	46	40.4	385	2	Q91YF8_MESAU	O91yf8 mesocricetu
87	46	40.4	385	2	Q91YH1_CRIMI	O91yh1 cricetus
88	46	40.4	410	2	Q841K7_9ACTO	O841k7 streptomyc
89	46	40.4	462	1	IL6RA_RAT	P22773 rattus norv
90	46	40.4	711	2	Q4M193_9BURK	Q4m193 burkholderi
91	46	40.4	2924	2	Q722C2_TOXGO	Q722c2 toxoplasma
92	45.5	39.9	567	2	Q63MT4_BURPS	Q63mt4 burkholderi
93	45.5	39.9	567	2	Q62DG8_BURMA	Q62dg8 burkholderi
94	45.5	39.9	1111	2	Q80YN4_RAT	O80yn4 rattus norv
95	45.5	39.9	1242	2	Q57U98_9TRYP	O57u99 trypanosoma
96	45	39.5	111	2	Q9X898_STRCO	Q9x898 streptomyc
97	45	39.5	140	2	Q6G6J6_STAAS	O6g6j6 staphylococ
98	45	39.5	140	2	Q6GDM1_STAAR	O6gdm1 staphylococ
99	45	39.5	140	2	Q5HD77_STAAC	O5hd77 staphylococ
100	45	39.5	140	2	Q7A3M1_STAAN	O7a3m3 staphylococ

ALIGNMENTS

```

RESULT 1
Q62MM4 HUMAN
ID Q62MM4_HUMAN PRELIMINARY; PRT; 435 AA.
AC Q62MM4
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16821.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=86118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Senba K., Nomura N., Miyajima N.,
RA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
[2]
RN NUCLEOTIDE SEQUENCE, AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2959974;
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Libermann T.A., Schlessinger J., Francke U.,
RA Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
[3]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS CYS-452; VAL-655 AND
RN ALA-1170.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE OF 1-191.
RX MEDLINE=87286898; PubMed=3039351;
RA Tal M., King C.R., Kraus M.H., Ullrich A., Schlessinger J., Givol D.;
RT "Human HER2 (neu) promoter: evidence for multiple mechanisms for
RT transcriptional initiation.";
RL Mol. Cell. Biol. 7:2597-2601(1987).
[5]
RN NUCLEOTIDE SEQUENCE OF 737-1031.
RX MEDLINE=86016729; PubMed=2959567;
RA Senba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-
RT erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
[6]
RN NUCLEOTIDE SEQUENCE OF 832-909.
RX TISSUE=Mammary carcinoma;
RA King C.R., Kraus M.H., Aaronson S.A.;
RT "Amplification of a novel v-erbB-related gene in a human mammary
RT carcinoma.";
RL Science 229:974-976(1985).
[7]
RN NUCLEOTIDE SEQUENCE OF 1081-1245, AND VARIANT ALA-1170.
RX MEDLINE=94000386; PubMed=8104414;
RA Sarkar F.H., Ball D.E., Li Y.W., Crissman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
[8]
RN IDENTIFICATION IN A COMPLEX WITH PIK3C2A AND EGFR, IDENTIFICATION IN A
RN COMPLEX WITH PIK3C2B AND EGFR, AND INTERACTION WITH PIK3C2B.
RX PubMed=10805725;
RA Arcaro A., Zvelebil M.J., Wallasch C., Ullrich A., Waterfield M.D.,
RA Domin J.;
RT "Class II phosphoinositide 3-kinases are downstream targets of
RT activated polyphosphate growth factor receptors.";
RL Mol. Cell. Biol. 20:3817-3830(2000).
[9]
RN INTERACTION WITH PLXNB1.
RX PubMed=15210733; DOI=10.1083/jcb.200312094;
RA Swiercz J.M., Kuner R., Offermanns S.;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 22:22:28 ; Search time 1834.09 Seconds  
(without alignments)  
619.854 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INTHSCVDLDDKGPAAQR 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09632036/runat\_02122005\_104140\_25811/app\_query.fasta\_1.398  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09632036 @CGN 1 1 3563 @runat\_02122005\_104140\_25811 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_ov:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_ets:\*  
11: gb\_ey:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	105	92.1	201	6	AR229723 Sequence
2	105	92.1	499	6	CS085267 Sequence
3	105	92.1	739	6	CS085266 Sequence

4	105	92.1	922	6	CS085259 Sequence
5	105	92.1	979	6	CS085265 Sequence
6	105	92.1	1219	6	CS085264 Sequence
7	105	92.1	1459	6	CS085263 Sequence
8	105	92.1	1699	6	CS085262 Sequence
9	105	92.1	1939	6	CS085261 Sequence
10	105	92.1	2028	6	CS020247 Sequence
11	105	92.1	2028	6	CS020248 Sequence
12	105	92.1	2083	6	CS085260 Sequence
13	105	92.1	2086	6	CS085268 Sequence
14	105	92.1	2086	6	CS085269 Sequence
15	105	92.1	2086	6	CS085270 Sequence
16	105	92.1	2086	6	CS085271 Sequence
17	105	92.1	2086	6	CS085272 Sequence
18	105	92.1	2320	6	CS075235 Sequence
19	105	92.1	2385	6	AR082744 Sequence
20	105	92.1	2385	6	AR099963 Sequence
21	105	92.1	2385	6	AR143949 Sequence
22	105	92.1	2437	6	CS075247 Sequence
23	105	92.1	2871	6	BD222937 Heteromin
24	105	92.1	2871	6	AX023363 Sequence
25	105	92.1	3678	6	AX050514 Sequence
26	105	92.1	3768	6	AR034479 Sequence
27	105	92.1	3768	6	BD224136 Novel met
28	105	92.1	3768	6	BD267514 HER-2/neu
29	105	92.1	3768	6	CQ815384 Sequence
30	105	92.1	3768	6	CQ841302 Sequence
31	105	92.1	3768	6	CQ841343 Sequence
32	105	92.1	3768	6	CQ841344 Sequence
33	105	92.1	3768	6	CS007994 Sequence
34	105	92.1	3768	6	CS020239 Sequence
35	105	92.1	3768	6	CS104222 Sequence
36	105	92.1	3768	6	AR409603 Sequence
37	105	92.1	3768	6	AX060704 Sequence
38	105	92.1	3768	6	AX201817 Sequence
39	105	92.1	3768	6	AX380923 Sequence
40	105	92.1	3768	6	AX384604 Sequence
41	105	92.1	3768	6	AX465456 Sequence
42	105	92.1	3768	6	AX467229 Sequence
43	105	92.1	3768	6	AX481438 Sequence
44	105	92.1	3768	11	AY888050 Synthetic
45	105	92.1	3768	11	AY888051 Synthetic
46	105	92.1	3768	11	AY888102 Synthetic
47	105	92.1	3768	11	AY888103 Synthetic
48	105	92.1	3768	11	AY888104 Synthetic
49	105	92.1	3778	6	CS020249 Sequence
50	105	92.1	4473	6	AR080259 Sequence
51	105	92.1	4473	6	AR167390 Sequence
52	105	92.1	4473	6	AR392088 Sequence
53	105	92.1	4473	6	AR534814 Sequence
54	105	92.1	4473	8	HSERB2R
55	105	92.1	4483	6	CQ723568 Sequence
56	105	92.1	4530	6	CQ834023 Sequence
57	105	92.1	4530	6	CQ873909 Sequence
58	105	92.1	4530	6	CQ977093 Sequence
59	105	92.1	4530	6	CS052366 Sequence
60	105	92.1	4530	6	CS063102 Sequence
61	105	92.1	4530	6	CS080883 Sequence
62	105	92.1	4530	6	CS101342 Sequence
63	105	92.1	4530	6	CS103534 Sequence
64	105	92.1	4530	6	I21124 Sequence 9
65	105	92.1	4530	6	I59745 Sequence 9
66	105	92.1	4530	6	AR202597 Sequence
67	105	92.1	4530	6	AR283481 Sequence
68	105	92.1	4530	6	AR344811 Sequence
69	105	92.1	4530	6	AX282577 Sequence
70	105	92.1	4530	6	AX587649 Sequence
71	105	92.1	4530	6	AX644071 Sequence
72	105	92.1	4530	6	AX771418 Sequence
73	105	92.1	4530	6	AX938565 Sequence
74	105	92.1	4530	6	BD005474 Cellular
75	105	92.1	4530	8	HUMHER2A
76	105	92.1	4536	8	BC080193 Homo sapi

77 105 92.1 4647 6 CS075243 Sequence  
78 105 92.1 4932 6 CQ853478 Sequence  
79 105 92.1 4932 8 AK131568 Homo sapi  
80 105 92.1 5661 6 CQ975755 Sequence  
81 105 92.1 9274 6 AR409602 Sequence  
82 105 92.1 9274 6 AX060703 Sequence  
83 100 87.7 3768 6 CQ841341 Sequence  
84 98 86.0 528 9 AF393158 Rattus no  
85 98 86.0 2763 6 AX380942 Sequence  
86 98 86.0 2781 6 AX380944 Sequence  
87 98 86.0 3771 6 BD267516 HER-2/neu  
88 98 86.0 3771 6 AX189662 Sequence  
89 98 86.0 3771 6 AX380925 Sequence  
90 98 86.0 3955 6 BD267515 Sequence  
91 98 86.0 3955 6 I21129 Sequence  
92 98 86.0 3955 6 I59750 Sequence  
93 98 86.0 3955 6 AX380924 Sequence  
94 98 86.0 3955 9 RNNEUR  
95 98 86.0 4062 9 HAMNEU  
96 98 86.0 4478 9 BC061863  
97 98 86.0 4506 9 AK129487 Rattus no  
98 98 86.0 4694 9 BC046811 Mus muscu  
99 98 86.0 4695 9 BC053078 Mus muscu  
100 98 86.0 4727 9 AY116182 Rattus no

## ALIGNMENTS

RESULT 1  
LOCUS AR229723 201 bp RNA linear PAT 20-DEC-2002  
DEFINITION Sequence 3 from patent US 6451524.  
ACCESSION AR229723  
VERSION AR229723.1 GI:27269547  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Ecker,D.J.  
TITLE Identification of disease predictive nucleic acids  
JOURNAL Patent: US 6451524-A 3 17-SEP-2002;  
ISIS Pharmaceuticals, Inc.; Carlsbad, CA  
FEATURES  
source  
1..201  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned RNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 6.97e-08 Length: 201  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-42 (1-20) x AR229723 (1-201)

Qy 1 IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 83 ATCAACTGCACCACTCCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 142

RESULT 2  
LOCUS CS085267 499 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 9 from Patent WO2005039618.  
ACCESSION CS085267  
VERSION CS085267.1 GI:66711224  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1

REFERENCE 1  
AUTHORS Amici,A., Cavallo,F., Forni,G. and Marchini,C.  
TITLE P185 sp neu /sp -encoding DNA and Therapeutical uses thereof  
JOURNAL Patent: WO 2005039618-A 9 06-MAY-2005;  
Indena S.p.A. (IT)  
FEATURES  
source  
1..499  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.63e-07 Length: 499  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-42 (1-20) x CS085267 (1-499)

Qy 1 IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 326 ATCAACTGCACCACTCCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 385

RESULT 3  
LOCUS CS085266 739 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 8 from Patent WO2005039618.  
ACCESSION CS085266  
VERSION CS085266.1 GI:66711223  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Amici,A., Cavallo,F., Forni,G. and Marchini,C.  
TITLE P185 sp neu /sp -encoding DNA and Therapeutical uses thereof  
JOURNAL Patent: WO 2005039618-A 8 06-MAY-2005;  
Indena S.p.A. (IT)  
FEATURES  
source  
1..739  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.35e-07 Length: 739  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-42 (1-20) x CS085266 (1-739)

Qy 1 IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 566 ATCAACTGCACCACTCCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 625

## ORIGIN

Alignment Scores:  
Pred. No.: 2.35e-07 Length: 739  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 6 Gaps: 0  
US-09-632-036F-42 (1-20) x CS085266 (1-739)

Qy 1 IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 566 ATCAACTGCACCACTCCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 625

## RESULT 4

LOCUS CS085259 922 bp DNA linear PAT 25-MAY-2005  
DEFINITION Sequence 1 from Patent WO2005039618.  
ACCESSION CS085259  
VERSION CS085259.1 GI:66711216  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 21:10:38 ; Search time 235.909 Seconds  
(without alignments)  
565.022 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSHSCVLDLDRGCPAEQR 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFTO\_spool/US09632036/runat\_02122005\_104140\_25801/app\_query.fasta\_1.398  
-DB=N\_Geneseq -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=100  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09632036 @CGN 1.1 714 @runat\_02122005\_104140\_25801 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_21.\*

1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002as.\*  
7: Geneseq2002bs.\*  
8: Geneseq2003as.\*  
9: Geneseq2003bs.\*  
10: Geneseq2003cs.\*  
11: Geneseq2003ds.\*  
12: Geneseq2004as.\*  
13: Geneseq2004bs.\*  
14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Match	Length	ID	Description
1	105	92.1	201	3	AAA62125	Aaa62125 Her2 norm
2	105	92.1	499	14	ADZ69878	Adz69878 p185neu o
3	105	92.1	655	10	ACD94035	Acd94035 Human col
4	105	92.1	739	14	ADZ69877	Adz69877 p185neu o

5	105	92.1	884	10	ADG17941	Adg17941 Human HER
6	105	92.1	922	14	ADZ69870	Adz69870 p185neu o
7	105	92.1	979	14	ADZ69876	Adz69876 p185neu o
8	105	92.1	1219	14	ADZ69875	Adz69875 p185neu o
9	105	92.1	1459	14	ADZ69874	Adz69874 p185neu o
10	105	92.1	1479	10	ADG17935	Adg17935 Human HER
11	105	92.1	1699	14	ADZ69873	Adz69873 p185neu o
12	105	92.1	1939	14	ADZ69872	Adz69872 p185neu o
13	105	92.1	1956	12	ADJ57171	Adj57171 Her-2/neu
14	105	92.1	2028	14	ADM87396	Adm87396 Human cod
15	105	92.1	2028	14	ADM87397	Adm87397 Human wil
16	105	92.1	2052	12	ADJ57170	Adj57170 Human HER
17	105	92.1	2083	14	ADZ69871	Adz69871 p185neu o
18	105	92.1	2086	14	ADZ69880	Adz69880 p185neu o
19	105	92.1	2086	14	ADZ69879	Adz69879 p185neu o
20	105	92.1	2086	14	ADZ69881	Adz69881 p185neu o
21	105	92.1	2086	14	ADZ69882	Adz69882 p185neu o
22	105	92.1	2086	14	ADZ69883	Adz69883 p185neu o
23	105	92.1	2125	2	AAV21727	Aav21727 Humanised
24	105	92.1	2149	10	ADG17943	Adg17943 Human HER
25	105	92.1	2164	10	ADM17939	Adm17939 Human HER
26	105	92.1	2261	11	ADM29355	Adm29355 Human nov
27	105	92.1	2320	14	ADZ47794	Adz47794 DNA encod
28	105	92.1	2385	2	AAT72725	Aat72725 Her2-GM-C
29	105	92.1	2437	14	ADZ47806	Adz47806 DNA encod
30	105	92.1	2802	13	ADT50728	Adt50728 Cancer re
31	105	92.1	2871	3	AZ50586	Aaz50586 DC8cFV-e
32	105	92.1	3188	11	ADM29351	Adm29351 Human nov
33	105	92.1	3600	3	AAA89736	Aaa89736 Human HER
34	105	92.1	3678	6	ABK86207	Abk86207 cDNA enco
35	105	92.1	3765	10	ADB67620	Adb67620 Human epi
36	105	92.1	3765	14	ADZ72379	Adz72379 DNA encod
37	105	92.1	3768	2	AAT40739	Aat40739 HER-2/neu
38	105	92.1	3768	2	AXA01912	Axa01912 Human HER
39	105	92.1	3768	3	AAA09455	Aaa09455 Human her
40	105	92.1	3768	5	AAH23392	Aah23392 Human HER
41	105	92.1	3768	6	ABV78168	Abv78168 Human ERB
42	105	92.1	3768	6	ABZ35744	Abz35744 Human ERB
43	105	92.1	3768	6	ABZ32743	Abz32743 Human HER
44	105	92.1	3768	6	ABK10730	Abk10730 Human HER
45	105	92.1	3768	6	ADA43935	Ada43935 Human HER
46	105	92.1	3768	6	ABX09987	Abx09987 Human ERB
47	105	92.1	3768	6	ADA43986	Ada43986 Human HER
48	105	92.1	3768	6	ABK14058	Abk14058 Human HER
49	105	92.1	3768	6	ABA92250	Ab92250 Human HER
50	105	92.1	3768	6	ABL91709	Ab191709 Human pol
51	105	92.1	3768	8	ACC57649	Acc57649 Human pro
52	105	92.1	3768	13	ADQ75970	Adq75970 Rhesus mo
53	105	92.1	3768	13	ADQ75929	Adq75929 Rhesus mo
54	105	92.1	3768	13	ADQ75971	Adq75971 Rhesus mo
55	105	92.1	3768	14	ADV43219	Adv43219 Human pay
56	105	92.1	3768	14	ADM87389	Adm87389 Human cod
57	105	92.1	3768	14	AEA47578	Aea47578 Nucleotid
58	105	92.1	3778	14	ADM87398	Adm87398 Human cod
59	105	92.1	4299	2	AAQ46083	Aaq46083 Sequence
60	105	92.1	4472	3	AAA14812	Aaa14812 cDNA enco
61	105	92.1	4473	2	ABO76220	Abq76220 Human tum
62	105	92.1	4473	2	AAZ31071	Aaz31071 HER-2 nuc
63	105	92.1	4473	6	ABZ34969	Abz34969 Human gen
64	105	92.1	4473	6	AAZ38904	Aaz38904 Human HER
65	105	92.1	4473	8	ACC69999	Acc69999 Human c-e
66	105	92.1	4473	10	ADC35148	Adc35148 Human bre
67	105	92.1	4473	10	ADD25483	Ad25483 Binding d
68	105	92.1	4473	12	AQJ17192	Adq17192 Human sof
69	105	92.1	4478	12	ADQ83974	Adq83974 Human tum
70	105	92.1	4529	12	ADJ57169	Adj57169 Human HER
71	105	92.1	4530	2	AAT01585	Aat01585 Her-2/neu
72	105	92.1	4530	2	AAT71253	Aat71253 Human HER
73	105	92.1	4530	3	AZ60815	Aaz60815 Nucleotid
74	105	92.1	4530	3	AAU19731	Aad19731 Human tyr
75	105	92.1	4530	6	ABN85585	Abn85585 Human HER
76	105	92.1	4530	6	ABZ35012	Abz35012 Human gen
77	105	92.1	4530	6	ABV94128	Abv94128 Breast ca





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 10, 2005, 22:40:44 ; Search time 78.1818 Seconds  
(without alignments)  
454.725 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSHSCVLDLDDKCPAEQR 20

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09632036/runat\_02122005\_104141\_25840/app\_query.fasta\_1.398  
-DB=Issued\_Patents\_NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTPMT=ptlo -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09632036 @CIGN 1.1.188 @runat\_02122005\_104141\_25840 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP.COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE.COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	201	3	US-09-200-355-3
2	105	92.1	2385	2	US-09-146-283-3
3	105	92.1	2385	3	US-08-579-823A-3
4	105	92.1	2385	3	US-09-344-195-3
5	105	92.1	3768	2	US-08-625-101-1
6	105	92.1	3768	2	US-08-356-786-1
7	105	92.1	3768	3	US-09-811-115-2
8	105	92.1	3768	1	US-09-167-516-1
9	105	92.1	4473	2	US-09-048-804-1

105	92.1	4473	3	US-09-056-105-26	Sequence 26, Appli	
105	92.1	4473	3	US-09-663-834A-3	Sequence 3, Appli	
105	92.1	4473	3	US-09-441-411-5	Sequence 5, Appli	
105	92.1	4530	2	US-08-229-515A-9	Sequence 9, Appli	
105	92.1	4530	2	US-08-645-865-9	Sequence 9, Appli	
105	92.1	4530	3	US-09-167-322-4	Sequence 4, Appli	
105	92.1	4530	3	US-09-527-487-1	Sequence 11, Appli	
105	92.1	4530	3	US-09-877-177A-11	Sequence 11, Appli	
105	92.1	9274	3	US-09-811-115-1	Sequence 1, Appli	
98	86.0	3955	2	US-08-229-515A-14	Sequence 14, Appli	
98	86.0	3955	2	US-08-645-865-14	Sequence 14, Appli	
85	74.6	1872	3	US-08-422-108-2	Sequence 2, Appli	
85	74.6	1872	3	US-08-422-73A-2	Sequence 2, Appli	
68	59.6	153	3	US-08-776-251-3	Sequence 3, Appli	
24	68	816	3	US-08-776-251-10	Sequence 10, Appli	
25	51.5	45.2	601	3	US-09-949-016-26918	Sequence 26918, A
26	51.5	45.2	601	3	US-09-949-016-26919	Sequence 26919, A
27	51.5	45.2	601	3	US-09-949-016-182960	Sequence 182960, A
28	51.5	45.2	601	3	US-09-949-016-182961	Sequence 182961, A
29	51.5	45.2	41737	3	US-09-949-016-12204	Sequence 12204, A
30	51.5	45.2	41741	3	US-09-949-016-16983	Sequence 16983, A
31	51.5	44.7	390890	3	US-09-949-016-14720	Sequence 14720, A
32	50	43.9	121982	3	US-09-949-016-12085	Sequence 12085, A
33	50	43.9	121982	3	US-09-949-016-14105	Sequence 14105, A
34	50	43.9	123025	3	US-09-198-452A-1	Sequence 1, Appli
35	50	43.9	1230230	3	US-09-438-185A-1	Sequence 1, Appli
36	48	42.1	264	2	US-08-313-185-65	Sequence 65, Appli
37	48	42.1	264	3	US-09-082-614A-65	Sequence 65, Appli
38	42.1	306	2	US-08-313-185-63	Sequence 63, Appli	
39	48	42.1	306	3	US-09-082-614A-63	Sequence 116779, A
40	48	42.1	601	3	US-09-949-016-116779	Sequence 116779, A
41	48	42.1	7790	3	US-09-949-016-16903	Sequence 16903, A
42	48	42.1	65561	3	US-09-949-016-15365	Sequence 15365, A
43	48	42.1	325034	3	US-09-949-016-14957	Sequence 14957, A
44	48	42.1	389504	3	US-09-949-016-11774	Sequence 11774, A
45	48	42.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
46	48	42.1	4411529	3	US-09-103-840A-2	Sequence 1, Appli
47	47.5	41.7	601	3	US-09-949-016-30625	Sequence 30625, A
47	47.5	41.7	601	3	US-09-949-016-73128	Sequence 73128, A
49	47.5	41.7	36171	3	US-09-949-016-13876	Sequence 13876, A
50	47.5	41.7	36274	3	US-09-949-016-12389	Sequence 12389, A
51	47	41.2	179	3	US-09-513-999C-29594	Sequence 29594, A
52	47	41.2	462	2	US-08-313-185-61	Sequence 61, Appli
53	47	41.2	462	3	US-09-082-614A-61	Sequence 61, Appli
54	47	41.2	1018	3	US-09-188-930-259	Sequence 259, App
55	47	41.2	1018	3	US-09-312-283C-259	Sequence 259, App
56	47	41.2	1531	2	US-08-833-963C-1	Sequence 1, Appli
57	47	41.2	1661	3	US-09-409-096-5	Sequence 5, Appli
58	47	41.2	1741	3	US-09-949-016-1921	Sequence 1921, Ap
59	47	41.2	2018	3	US-08-980-514-2	Sequence 2, Appli
60	47	41.2	2736	3	US-09-220-132-66	Sequence 66, Appli
61	47	41.2	31867	3	US-09-949-016-16874	Sequence 16874, A
62	47	41.2	31875	3	US-09-949-016-12864	Sequence 12864, A
63	47	41.2	119981	3	US-09-949-016-11844	Sequence 11844, A
64	47	41.2	119982	3	US-09-949-016-13606	Sequence 13606, A
65	46	40.4	400	3	US-08-956-171E-3629	Sequence 3629, Ap
66	46	40.4	400	3	US-08-781-986A-3629	Sequence 3629, Ap
67	46	40.4	601	3	US-09-949-016-60579	Sequence 60579, A
68	46	40.4	601	3	US-09-949-016-60580	Sequence 60580, A
69	46	40.4	601	3	US-09-949-016-131286	Sequence 131286, A
70	46	40.4	2334	3	US-08-936-165A-206	Sequence 206, App
71	46	40.4	4330	3	US-09-310-293-1	Sequence 1, Appli
72	46	40.4	4330	3	US-09-579-376-1	Sequence 1, Appli
73	46	40.4	5238	3	US-09-620-312D-351	Sequence 351, App
74	46	40.4	14051	3	US-08-956-171E-103	Sequence 103, App
75	46	40.4	14051	3	US-08-781-986A-103	Sequence 103, App
76	46	40.4	35675	3	US-09-949-016-13505	Sequence 13505, A
77	46	40.4	40781	3	US-09-949-016-17016	Sequence 17016, A
78	46	40.4	199945	3	US-09-949-016-15436	Sequence 15436, A
79	46	40.4	24781	3	US-09-949-016-14193	Sequence 14193, A
80	45.5	39.9	2886	3	US-09-694-777A-13	Sequence 13, Appli
81	45.5	39.9	2967	3	US-09-694-777A-14	Sequence 14, Appli
82	45.5	39.9	3002	3	US-09-694-777A-1	Sequence 1, Appli

83 45.5 39.9 3083 3 US-09-694-777A-2 Sequence 2, Appli  
84 45 39.5 374 3 US-09-270-767-40 Sequence 40, Appl  
85 45 39.5 374 3 US-09-270-767-15322 Sequence 15322, A  
86 45 39.5 519 3 US-09-621-976-10460 Sequence 10460, A  
87 45 39.5 531 4 US-09-605-703B-2597 Sequence 2597, Ap  
88 45 39.5 612 3 US-09-248-796A-1530 Sequence 1530, Ap  
89 45 39.5 700 3 US-09-735-271-61 Sequence 61, Appl  
90 45 39.5 737 3 US-09-533-559-4539 Sequence 4539, Ap  
91 45 39.5 1029 2 US-09-902-540-8675 Sequence 8675, Ap  
92 45 39.5 1108 2 US-07-866-560-5 Sequence 5, Appli  
93 45 39.5 1108 2 US-08-077-673-5 Sequence 5, Appli  
94 45 39.5 1108 2 US-08-478-992-5 Sequence 5, Appli  
95 45 39.5 1108 3 US-09-105-298-5 Sequence 5, Appli  
96 45 39.5 1108 3 US-08-706-281A-9 Sequence 9, Appli  
97 45 39.5 1108 3 US-09-097-231-9 Sequence 9, Appli  
98 45 39.5 1108 3 US-09-542-122-5 Sequence 5, Appli  
99 45 39.5 1108 3 US-09-353-099-9 Sequence 9, Appli  
100 45 1631 3 US-09-149-476-97 Sequence 97, Appl

## ALIGNMENTS

RESULT 1  
US-09-200-355-3  
; Sequence 3, Application US/09200355  
; Patent No. 6451524  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/09/200,355  
; CURRENT FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-200-355-3

Alignment Scores:  
Pred. No.: 2.2e-09 Length: 201  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 3 Gaps: 0

US-09-632-036F-42 (1-20) x US-09-200-355-3 (1-201)

Qy 1 ileAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 83 AUCACUGCACCACUCCUGUGUGGACCUUGAUGACAGGGGCGCCGCCGAGCAGAGA 142

## RESULT 2

US-09-146-283-3  
; Sequence 3, Application US/09146283  
; Patent No. 5976546  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Compositions  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94106  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,283  
; FILING DATE: 03-SEPT-1998  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2385 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: homo sapiens  
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8  
US-09-146-283-3

Alignment Scores:  
Pred. No.: 4.13e-08 Length: 2385  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 2 Gaps: 0

US-09-632-036F-42 (1-20) x US-09-146-283-3 (1-2385)

Qy 1 ileAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 1892 ATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAGGGGTGCCCGCCGAGCAGAGA 1951

## RESULT 3

US-08-579-823A-3  
; Sequence 3, Application US/08579823A  
; Patent No. 6080409  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Composition and Method  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94106  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/579,823A  
; FILING DATE: 03-DEC-1998  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010  
; TELECOMMUNICATION INFORMATION:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 11, 2005, 00:00:23 ; Search time 391.818 Seconds  
(without alignments)  
422.103 Million cell updates/sec

Title: US-09-632-036F-42

Perfect score: 114

Sequence: 1 INGTSCVDLDKCPAEQR 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USFTO\_spool/US09632036/runat\_02122005\_104143\_25924/app\_query.fasta\_1.398  
-DB=Published Applications NA\_Main -QWMT=fastcap -SUFFIX=rnpbm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORMS=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09632036 cGCG 1.1134 @runat\_02122005\_104143\_25924  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOC -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	201	5	US-10-109-213-3
2	105	92.1	884	6	US-10-412-804A-7
3	105	92.1	1479	6	US-10-412-804A-1
4	105	92.1	2149	6	US-10-412-804A-9
5	105	92.1	2164	6	US-10-412-804A-5
6	105	92.1	2320	9	US-10-956-373-1
7	105	92.1	2437	9	US-10-956-373-13
8	105	92.1	3504	9	US-10-794-514A-329
Sequence 329, App					
Sequence 3, Appli					
Sequence 7, Appli					
Sequence 1, Appli					
Sequence 9, Appli					
Sequence 5, Appli					
Sequence 1, Appli					
Sequence 139, App					

9	105	92.1	3765	6	US-10-207-498-5	Sequence 5, Appli
10	105	92.1	3768	3	US-09-811-123-8	Sequence 8, Appli
11	105	92.1	3768	3	US-09-811-115-2	Sequence 2, Appli
12	105	92.1	3768	3	US-09-765-973-1	Sequence 1, Appli
13	105	92.1	3768	3	US-09-854-356-9	Sequence 9, Appli
14	105	92.1	3768	3	US-09-930-125-1	Sequence 1, Appli
15	105	92.1	3768	3	US-09-984-092-3	Sequence 3, Appli
16	105	92.1	3768	6	US-10-313-644-1	Sequence 1, Appli
17	105	92.1	3768	7	US-10-280-576-3	Sequence 3, Appli
18	105	92.1	3768	7	US-10-441-799C-3	Sequence 3, Appli
19	105	92.1	3768	7	US-10-384-339C-52	Sequence 52, Appli
20	105	92.1	4473	3	US-09-441-411-5	Sequence 5, Appli
21	105	92.1	4473	3	US-10-146-473-32	Sequence 32, Appli
22	105	92.1	4473	5	US-10-207-655-44	Sequence 44, Appli
23	105	92.1	4473	6	US-10-101-510-81	Sequence 81, Appli
24	105	92.1	4473	8	US-10-762-128-5	Sequence 5, Appli
25	105	92.1	4473	8	US-10-723-860-8	Sequence 8, Appli
26	105	92.1	4530	3	US-09-877-177-11	Sequence 11, Appli
27	105	92.1	4530	6	US-10-177-923-125	Sequence 125, App
28	105	92.1	4530	6	US-10-007-926A-119	Sequence 119, App
29	105	92.1	4530	6	US-10-338-730-1	Sequence 1, Appli
30	105	92.1	4530	6	US-10-101-510-124	Sequence 124, App
31	105	92.1	4530	6	US-10-116-275-131	Sequence 131, App
32	105	92.1	4530	6	US-10-426-836-11	Sequence 11, Appli
33	105	92.1	4530	6	US-10-272-437A-27	Sequence 27, Appli
34	105	92.1	4530	6	US-10-117-937-595	Sequence 595, App
35	105	92.1	4530	6	US-10-392-113-45	Sequence 45, Appli
36	105	92.1	4530	6	US-10-159-563-208	Sequence 208, App
37	105	92.1	4530	7	US-10-435-696-10	Sequence 10, Appli
38	105	92.1	4530	7	US-10-734-564-59	Sequence 59, Appli
39	105	92.1	4530	8	US-10-657-022-91	Sequence 91, Appli
40	105	92.1	4530	8	US-10-615-343-16	Sequence 16, Appli
41	105	92.1	4530	9	US-10-928-465-75	Sequence 75, Appli
42	105	92.1	4530	9	US-10-794-514A-2	Sequence 2, Appli
43	105	92.1	4530	9	US-10-871-708-18	Sequence 18, Appli
44	105	92.1	4530	10	US-11-067-064-595	Sequence 595, App
45	105	92.1	4530	10	US-11-067-159-595	Sequence 595, App
46	105	92.1	4543	3	US-09-769-508-1	Sequence 1, Appli
47	105	92.1	4606	3	US-09-971-392-70	Sequence 70, Appli
48	105	92.1	4642	5	US-10-198-846-10896	Sequence 10896, A
49	105	92.1	4647	9	US-10-956-373-9	Sequence 9, Appli
50	105	92.1	5125	8	US-10-723-860-4619	Sequence 4619, Ap
51	105	92.1	9274	3	US-09-811-123-7	Sequence 7, Appli
52	105	92.1	9274	3	US-09-811-115-1	Sequence 1, Appli
53	98	86.0	3771	3	US-09-854-356-11	Sequence 11, Appli
54	98	86.0	3952	3	US-10-949-667-9	Sequence 9, Appli
55	98	86.0	3955	3	US-09-870-759-117	Sequence 117, App
56	98	86.0	3955	3	US-09-854-356-10	Sequence 10, Appli
57	98	86.0	3955	3	US-09-751-708A-117	Sequence 117, App
58	98	86.0	3955	8	US-10-428-817A-113	Sequence 113, App
59	98	86.0	3955	9	US-10-937-758A-94	Sequence 94, Appli
60	88.5	77.6	3564	9	US-10-794-514A-331	Sequence 331, App
61	82	71.9	825	8	US-10-484-067-9	Sequence 9, Appli
62	81	71.1	151	8	US-10-484-067-12	Sequence 12, Appli
63	78	68.4	154	8	US-10-484-067-6	Sequence 6, Appli
64	78	68.4	828	8	US-10-484-067-3	Sequence 3, Appli
65	57	50.0	1167	7	US-10-282-122A-9291	Sequence 9291, Ap
66	56	49.1	1941	6	US-10-108-260A-2086	Sequence 2086, Ap
67	55	48.2	589	4	US-09-925-065A-537341	Sequence 537341, A
68	53	46.5	334	7	US-10-424-599-51731	Sequence 51731, A
69	52.5	46.1	2295	9	US-10-114-270-107	Sequence 107, App
70	52.5	46.1	2597	9	US-10-450-763-10485	Sequence 10485, A
71	52.5	46.1	2674	9	US-10-956-157-2120	Sequence 2120, Ap
72	52	45.6	778	5	US-10-027-632-162468	Sequence 162468, A
73	52	45.6	778	6	US-10-027-632-162468	Sequence 162468, A
74	52	45.6	1467	10	US-11-097-143-13199	Sequence 13199, A
75	52	45.6	2846	10	US-11-097-143-13198	Sequence 13198, A
76	52	45.6	4584	10	US-11-097-143-13198	Sequence 13198, A
77	51.5	45.2	1223197	5	US-10-027-632-179264	Sequence 179264, A
78	51.5	45.2	1223197	6	US-10-027-632-179264	Sequence 179264, A
79	51	44.7	602	4	US-09-925-065A-695694	Sequence 695694, A
80	51	44.7	1064	4	US-09-925-065A-61266	Sequence 61266, A
81	51	44.7	97247	5	US-10-087-192-688	Sequence 688, App

```
C 82      51      44.7 518360 7 US-10-367-094-125      Sequence 125, App
C 83      50.5      44.3 956 8 US-10-774-355A-824      Sequence 824, App
C 84      50.5      44.3 1915 4 US-09-925-065A-78681      Sequence 78681, A
C 85      50      43.9 349 8 US-10-425-115-123077      Sequence 123077,
C 86      50      43.9 369 7 US-10-282-122A-18540      Sequence 18540, A
C 87      50      43.9 369 7 US-10-282-122A-19114      Sequence 19114, A
C 88      50      43.9 1783 7 US-10-425-114-15485      Sequence 15485, A
C 89      50      43.9 2568 8 US-10-491-213-77      Sequence 77, Appl
C 90      50      43.9 2821 8 US-10-425-115-46108      Sequence 46108, A
C 91      50      43.9 4766 3 US-09-764-868-1443      Sequence 1443, Ap
C 92      50      43.9 15425 3 US-09-764-869-1654      Sequence 1654, Ap
C 93      50      43.9 15425 5 US-10-091-504-1654      Sequence 1654, Ap
C 94      50      43.9 15425 6 US-10-227-577-1654      Sequence 1654, Ap
C 95      50      43.9 26332 5 US-10-087-192-1558      Sequence 1558, Ap
C 96      50      43.9 43623 8 US-10-741-600-17818      Sequence 17818, A
C 97      50      43.9 111331 9 US-10-461-862-101      Sequence 101, App
C 98      50      43.9 119501 6 US-10-174-319-15      Sequence 15, Appl
C 99      50      43.9 1230025 6 US-10-289-762-1      Sequence 1, Appli
C 100      49.5      43.4 553 5 US-10-027-632-266944      Sequence 266944,
```

## ALIGNMENTS

```
RESULT 1
US-10-109-213-3
; Sequence 3, Application US/10109213
; Publication No. US20020168670A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids
; FILE REFERENCE: IBIS0009
; CURRENT APPLICATION NUMBER: US/10/109,213
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US/09/200,355
; PRIOR FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-109-213-3

Alignment Scores:
Pred. No.:      3.13e-09      Length:      201
Score:          105.00      Matches:      19
Percent Similarity: 95.00%      Conservative: 0
Best Local Similarity: 95.00%      Mismatches: 1
Query Match:      92.11%      Indels:      0
DB:              5      Gaps:      0

US-09-632-036F-42 (1-20) x US-10-109-213-3 (1-201)

Qy      1      IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20
      |||||
Db      83      AUCAAUGCACCACCCUGUGGACCUUGAGCAAGGGCTGCCCGCCGAGCAGAGA 142

RESULT 2
US-10-412-804A-7
; Sequence 7, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0

Qy      1      IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20
      |||||
Db      1321      ATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 1380

ALIGNMENTS

US-10-412-804A-1
; Sequence 1, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0

Qy      1      IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20
      |||||
Db      633      ATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 692

RESULT 3
US-10-412-804A-1
; Sequence 1, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1169)
; OTHER INFORMATION: "n" can be a, g, c, or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1170)
; OTHER INFORMATION: "n" can be a, g, c, or t
US-10-412-804A-1

Alignment Scores:
Pred. No.:      2.86e-08      Length:      1479
Score:          105.00      Matches:      19
Percent Similarity: 95.00%      Conservative: 0
Best Local Similarity: 95.00%      Mismatches: 1
Query Match:      92.11%      Indels:      0
DB:              6      Gaps:      0

US-09-632-036F-42 (1-20) x US-10-412-804A-1 (1-1479)

Qy      1      IleAsnGlyThrHisSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20
      |||||
Db      1321      ATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAGGGCTGCCCGCCGAGCAGAGA 1380

RESULT 4
US-10-412-804A-9
; Sequence 9, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 11, 2005, 01:00:20 ; Search time 119.545 Seconds  
(without alignment)  
62.546 Million cell updates/sec

Title: US-09-632-036F-42  
Perfect score: 114  
Sequence: 1 INGTSCVDLDRGCPAEQR 20

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

## Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2_1/USPFO_spool/US09632036/runat_02122005_104143_25959/app_query.fasta_1.398  
-DB=Published Applications NA New -QFMT=fastcap -SUFFIX=rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62  
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USERS=US09632036 -CGN 1 1 227 @runat_02122005_104143_25959  
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

## Databases :

```
Published Applications NA New:  
1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:  
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:  
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:  
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:  
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:  
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:  
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:  
8: /cgn2_6/ptodata/2/pubpna/US12_NEW_PUB.seq:  
9: /cgn2_6/ptodata/2/pubpna/US13_NEW_PUB.seq:  
10: /cgn2_6/ptodata/2/pubpna/US14_NEW_PUB.seq:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	92.1	4530	US-10-770-726-18	Sequence 18, Appl
2	85	74.6	149419	US-11-112-908-49	Sequence 49, Appl
3	85	74.6	157224	US-11-112-908-51	Sequence 51, Appl
4	85	74.6	161726	US-11-112-908-48	Sequence 48, Appl
5	85	74.6	161726	US-11-112-908-52	Sequence 52, Appl
6	54	47.4	159497	US-11-112-908-61	Sequence 61, Appl
7	54	47.4	171427	US-11-112-908-60	Sequence 60, Appl
8	48	42.1	882	US-10-750-185-39806	Sequence 39806, A

47	41.2	784	7	US-11-112-908-418	Sequence 418, App
47	41.2	1178	6	US-10-750-185-49750	Sequence 49750, A
47	41.2	1875	6	US-10-131-826A-317	Sequence 317, App
47	41.2	177623	7	US-11-112-908-41	Sequence 41, Appl
46	40.4	693	6	US-10-750-185-52252	Sequence 52252, A
13	40.4	175673	7	US-11-121-086-55	Sequence 55, Appl
15	39.5	966	6	US-10-750-185-48688	Sequence 48688, A
45	39.5	3109	6	US-10-689-742-85	Sequence 85, Appl
45	39.5	172111	7	US-11-121-086-28	Sequence 28, Appl
45	39.5	189539	7	US-11-121-086-16	Sequence 16, Appl
19	44.5	1450	7	US-11-184-574-1	Sequence 5, Appl
20	44.5	3946	7	US-11-184-574-1	Sequence 1, Appl
21	44	2046	6	US-10-750-185-50776	Sequence 50776, A
22	44	2089	6	US-10-750-185-36037	Sequence 36037, A
23	44	169495	7	US-11-121-086-61	Sequence 61, Appl
24	43.5	1718	6	US-10-750-185-42547	Sequence 42547, A
25	43	37.7	1682	US-10-750-185-61950	Sequence 61950, A
26	43	37.7	1928	US-10-750-185-42585	Sequence 42585, A
27	43	37.7	2586	US-10-750-185-24808	Sequence 24808, A
28	43	37.7	2653	US-10-750-185-58741	Sequence 58741, A
29	43	37.7	3960	US-10-750-185-33376	Sequence 33376, A
30	43	37.7	4441	US-10-750-185-54843	Sequence 54843, A
31	43	37.7	7762	US-10-750-185-44139	Sequence 44139, A
32	43	37.7	110847	US-11-121-086-11	Sequence 11, Appl
33	43	37.7	153376	US-11-121-086-5	Sequence 5, Appl
34	43	37.7	1082144	US-11-117-187-211	Sequence 211, App
35	42.5	37.3	158	US-11-060-659-30	Sequence 30, Appl
36	42.5	37.3	159	US-11-060-659-31	Sequence 31, Appl
37	42.5	37.3	606	US-10-793-626-1513	Sequence 1513, Ap
38	42.5	37.3	2694	US-11-137-465-2	Sequence 2, Appl
39	42.5	37.3	2982	US-11-137-465-3	Sequence 3, Appl
40	42.5	37.3	3852	US-10-793-626-3374	Sequence 3374, Ap
41	42.5	37.3	4805	US-10-793-626-3382	Sequence 3382, Ap
42	42	36.8	600	US-10-750-185-1615	Sequence 1615, Ap
43	42	36.8	1120	US-10-750-185-43200	Sequence 43200, A
44	42	36.8	1195	US-10-750-185-57145	Sequence 57145, A
45	42	36.8	1610	US-10-750-185-45610	Sequence 45610, A
46	42	36.8	1671	US-10-750-185-56074	Sequence 56074, A
47	42	36.8	1689	US-10-750-185-43069	Sequence 43069, A
48	42	36.8	2081	US-11-186-284-135	Sequence 135, App
49	42	36.8	2105	US-11-186-284-133	Sequence 133, App
50	42	36.8	2356	US-10-750-185-46730	Sequence 46730, A
51	42	36.8	2356	US-10-750-185-61747	Sequence 61747, A
52	42	36.8	2513	US-10-750-185-38933	Sequence 38933, A
53	42	36.8	2598	US-11-102-240-151	Sequence 151, App
54	42	36.8	2767	US-10-821-234-43	Sequence 43, Appl
55	42	36.8	3398	US-10-750-185-64179	Sequence 64179, A
56	42	36.8	4545	US-10-750-185-37816	Sequence 37816, A
57	42	36.8	9454	US-11-109-056-1	Sequence 1, Appl
58	42	36.8	162537	US-11-121-086-59	Sequence 59, Appl
59	42	36.8	200828	US-11-121-086-62	Sequence 62, Appl
60	41.5	693	6	US-10-750-185-31742	Sequence 31742, A
61	41.5	36.4	877	US-10-750-185-37922	Sequence 37922, A
62	41.5	36.4	1615	US-10-750-185-56545	Sequence 56545, A
63	41.5	36.4	2176	US-10-750-185-32728	Sequence 32728, A
64	41.5	36.4	2535	US-10-750-185-50255	Sequence 50255, A
65	41	36.0	21	US-10-770-726-19133	Sequence 19133, A
66	41	36.0	938	US-10-750-185-43395	Sequence 43395, A
67	41	36.0	974	US-10-750-185-31530	Sequence 31530, A
68	41	36.0	996	US-10-750-185-54605	Sequence 54605, A
69	41	36.0	1038	US-10-750-185-38658	Sequence 38658, A
70	41	36.0	1134	US-10-750-185-59181	Sequence 59181, A
71	41	36.0	1438	US-10-750-185-43108	Sequence 43108, A
72	41	36.0	1438	US-10-750-185-45769	Sequence 45769, A
73	41	36.0	1700	US-10-750-185-63378	Sequence 63378, A
74	41	36.0	2228	US-10-750-185-61216	Sequence 61216, A
75	41	36.0	2244	US-10-750-185-35975	Sequence 35975, A
76	41	36.0	2373	US-10-750-185-55608	Sequence 55608, A
77	41	36.0	3024	US-10-750-185-26035	Sequence 26035, A
78	41	36.0	3391	US-10-750-185-64852	Sequence 64852, A
79	41	36.0	131855	US-11-112-908-29	Sequence 29, Appl
80	41	36.0	143389	US-11-112-908-30	Sequence 30, Appl
81	41	36.0	150173	US-11-112-908-26	Sequence 26, Appl

Sequence 38, Appl  
Sequence 30, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 103, Appl  
Sequence 28, Appl  
Sequence 27, Appl  
Sequence 24, Appl  
Sequence 102, Appl  
Sequence 15, Appl  
Sequence 1370, Appl  
Sequence 21, Appl  
Sequence 786, Appl  
Sequence 2013, Appl  
Sequence 43622, A  
Sequence 55567, A  
Sequence 47104, A  
Sequence 63683, A  
Sequence 32911, A

Sequence 38, Appl  
Sequence 30, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 103, Appl  
Sequence 28, Appl  
Sequence 27, Appl  
Sequence 24, Appl  
Sequence 102, Appl  
Sequence 15, Appl  
Sequence 1370, Appl  
Sequence 21, Appl  
Sequence 786, Appl  
Sequence 2013, Appl  
Sequence 43622, A  
Sequence 55567, A  
Sequence 47104, A  
Sequence 63683, A  
Sequence 32911, A

## ALIGNMENTS

RESULT 1  
US-10-770-726-18  
; Sequence 18, Application US/10770726  
; Publication No. US20050266409A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Brown, Eugene  
; APPLICANT: Liu, Wei  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; FILE REFERENCE: AM101079 (031896-010000)  
; CURRENT APPLICATION NUMBER: US/10770,726  
; CURRENT FILING DATE: 2004-02-04  
; NUMBER OF SEQ ID NOS: 48640  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-770-726-18

Alignment Scores:  
Pred. No.: 2.02e-07 Length: 4530  
Score: 105.00 Matches: 19  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 95.00% Mismatches: 1  
Query Match: 92.11% Indels: 0  
DB: 6 Gaps: 0

US-09-632-036F-42 (1-20) x US-10-770-726-18 (1-4530)  
Qy 1 ILAeSnGlyThriSerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 2032 ATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCCGCCGACGACGAGA 2091

RESULT 2  
US-11-112-908-49  
; Sequence 49, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01

Alignment Scores:  
Pred. No.: 0.0287 Length: 157224  
Score: 85.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 74.56% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-42 (1-20) x US-11-112-908-49 (1-149419)  
Qy 6 SerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 13921 AGCTGTGTGGACCTGGATGACAAGGGCTGCCCGCCGACGACGAGA 13965

RESULT 3  
US-11-112-908-51  
; Sequence 51, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 51  
; LENGTH: 157224  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-51

Alignment Scores:  
Pred. No.: 0.0287 Length: 157224  
Score: 85.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 74.56% Indels: 0  
DB: 7 Gaps: 0

US-09-632-036F-42 (1-20) x US-11-112-908-51 (1-157224)  
Qy 6 SerCysValAspLeuAspAspLysGlyCysProAlaGluGlnArg 20  
Db 150398 AGCTGTGTGGACCTGGATGACAAGGGCTGCCCGCCGACGACGAGA 150442

RESULT 4  
US-11-112-908-48  
; Sequence 48, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole